



PÁZMÁNY

Pázmány Péter Katolikus Egyetem
Jog- és Államtudományi Kar

J
A
S

iustum
aequum
salutare

Jogtudományi folyóirat

Budapest

XXI. 2025/3.

A szerkesztőbizottság társelnökei:

Burián László professor emeritus (PPKE JÁK)
Kukorelli István professor emeritus (ELTE ÁJK)

Főszerkesztő: Koltay András egyetemi tanár (PPKE JÁK)

Szerkesztőség:

Csink Lóránt egyetemi tanár (PPKE JÁK)
Koltay András egyetemi tanár (PPKE JÁK)
Szabó Sarolta egyetemi docens (PPKE JÁK)

A szerkesztőbizottság tagjai:

Bándi Gyula egyetemi tanár (PPKE JÁK)
Burián László professor emeritus (PPKE JÁK)
Csehi Zoltán egyetemi tanár (PPKE JÁK)
Harsági Viktória egyetemi tanár (PPKE JÁK)
Horváth Attila egyetemi tanár (NKE ÁNTK)
Kecskés László egyetemi tanár (PTE ÁJK)
Király Miklós egyetemi tanár (ELTE ÁJK)
Kovács Péter egyetemi tanár (PPKE JÁK)
Kukorelli István professor emeritus (ELTE ÁJK)
Mezey Barna egyetemi tanár (ELTE ÁJK)
Révész T. Mihály c. egyetemi tanár (SZIE DF ÁJK)
Schanda Balázs egyetemi tanár (PPKE JÁK)
Trócsányi László egyetemi tanár (KRE ÁJK)
Varga Zs. András egyetemi tanár (PPKE JÁK)
Várnay Ernő egyetemi tanár (DE ÁJK)
Vékás Lajos professor emeritus (ELTE ÁJK)

Technikai szerkesztő és titkár: Szakaliné Szeder Andrea

A szerkesztőség elérhetőségei:

Pázmány Péter Katolikus Egyetem Jog- és Államtudományi Kar
Dékáni Titkárság
1088 Budapest, Szentkirályi u. 28–30.
Telefon: (+36) 20 239 4247
www.ias.jak.ppke.hu
E-mail: ias@jak.ppke.hu

Kiadja:

a Pázmány Péter Katolikus Egyetem
Budapest, Szentkirályi u. 28–30.
www.ppke.hu

Nyomdai munkák: PPKE Egyetemi Nyomda

A folyóirat tartalma anonim módon lektorált.

TARTALOM

ACTA

Hollósi Gábor Első emléktörvényeink nyomában	5
Manzinger Krisztián Küzdelem a lett államnyelvért.	21

DISSERTATIONES

Bevezető a mesterséges intelligencia és a média kapcsolatát tárgyaló tanulmányokhoz.	41
Andok Mónika Mesterségesintelligencia-jártasság	43
Gosztonyi Gergely – Gyetván Dorina – Kovács Andrea A mesterséges intelligencia szólásszabadsága és tartalommoderálási gyakorlata a közösségi médiában a DSA és az MI rendelet vonatkozásában	61
Hohmann Balázs Átláthatóságot biztosító szabályok és intézkedések az MI tartalommoderálási tevékenysége során – kiemelt figyelemmel a DSA követelményrendszerére.	79
Lendvai Gergely Ferenc Algoritmikus elfogultság mint jogi dilemma	97
Mezei Kitti Algoritmusok és átláthatóság. Az online platformok ajánlórendszereinek jogi szabályozása az Európai Unióban	115
Pintér Róbert Új korszakot nyit-e a mesterséges intelligencia a dezinformációban? Az MI szerepének értelmezése a dezinformációs lánc, mint új elméleti keret segítségével.	135

Pongó Tamás
Útmutató egy iskolai (cyber)bullying szabályzat megalkotásához.161

Üveges István
A mesterséges intelligencia alapú tartalomszűrő algoritmusok társadalmi
hatásai, különös tekintettel a közösségimédia-platformokra177

VARIA

Beszámoló

VI. Fogyasztóvédelmi aktualitások – a digitalizáció,
a mesterséges intelligencia fogyasztóvédelmi kihívásai
(Veres Zoltán – Hámori Antal)197

SUMMA203

ACTA

ELSŐ EMLÉKTÖRVÉNYEINK NYOMÁBAN

Hollósi Gábor*

1. Bevezető

Törvényeink és az azokhoz kapcsolódó országgyűlési (nemzetgyűlési) anyagok történeti tartalmának vizsgálata évek óta foglalkoztat, legelőször ilyen szempontból az 1920. évi I. törvénycikk indokolását tanulmányoztam.¹ Ez a munkám volt az előzménye annak, hogy végül becsatlakoztam a VERITAS Történetkutató Intézet és Levéltár *emlékezettörténeti* kutatási programjába, melyben úgy a történész, mint a jogász szakma képviselőként a hazai *emléktörvények* monografikus feldolgozására vállalkoztam. Elsőként – főleg a két háború közötti időszak kutatójaként – a Horthy-korszak emléktörvényeiről szóló összefoglalásomat publikáltam a projekt addigi eredményeit összegző legelső kötetben,² az így elindult sorozatban pedig a második elemzésem „Emléktörvényeink a népi demokráciától a rendszerváltoztatásig (1945–1989)” címmel jelent meg.³ A dualizmus időszakának emléktörvényeit feldolgozó munkám egyelőre még kéziratban van, csak egy nagyon tömör összefoglalást közöltem belőle pusztán ismeretterjesztő szándékkal az *Élet és Tudomány* című folyóiratban.⁴ Jelenleg az 1990 és 2025 között keletkezett emléktörvények feldolgozásával foglalkozom; kismonográfiám megjelenése, mely az „Emléktörvényeink” címet fogja viselni, reményeim szerint 2027-ben lesz várható.

* Habilitált tudományos főmunkatárs, VERITAS Történetkutató Intézet és Levéltár.
ORCID: <https://orcid.org/0009-0009-3306-5728>

¹ Hollósi Gábor: A történelmi emlékezet az 1920. évi I. törvénycikk indokolásában. *Jogtörténeti Szemle*, 2019/1. 40–45.

² Hollósi Gábor: A Horthy-korszak emléktörvényei. In: Kincses Katalin Mária (szerk.): *XIX–XX. századi magyar történelmi kultuszok*. Budapest, Line Design, 2024. 141–174.

³ Hollósi Gábor: Emléktörvényeink a népi demokráciától a rendszerváltoztatásig (1945–1989). In: Kincses Katalin Mária (szerk.): *XIX–XX. századi magyar történelmi kultuszok II*. Budapest, Line Design, 2025. 211–255.

⁴ Hollósi Gábor: Történelem a törvénykönyv lapjain. A dualizmus korának emléktörvényei. *Élet és Tudomány*, 2025/10. 298–301.



Készülő könyvem első fejezetének tervezett anyagát mutatom be ezen a helyen, melyben egyrészt a legelső emléktörvényünket keresem, másrészt csoportosítását adom az általam korai jelzővel ellátott időszak hasonló jogalkotási termékeinek. Néhány, témámat érintő szakirodalmi megállapításra e tanulmány zárófejezetében kívánom kifejtetni az álláspontomat. Vizsgálódásom megkezdése előtt azonban elengedhetetlen, hogy tömören ismertessem azt a fogalmi keretet, amelyben a korai emléktörvényeinket keresem.

2. Fogalomhasználat

Az emléktörvény fogalmának meghatározására az utóbbi években magyar kutatók is kiemelt figyelmet fordítottak, elég utalnom Bán Marinának 2020-ban az amszterdami egyetemen,⁵ illetve Izsák-Somogyi Katalinnak 2024-ben a Pázmány Péter Katolikus Egyetemen megvédett doktori munkájára.⁶ Az utóbbi szerző magyar nyelven írt értekezése a hazai tudomány számára rendkívül értékes, hiszen szintézisét adja egy egész sor külföldi szerző – mint R. Badinter, U. Belavusau, F. Chandernagor, A. De Baets, C. Fortier, E. Fronza, S. Garibian, A. Gliszczynska-Grabias, E. Heinze, N. Koposov, F. Krawatzek, M. L. Lerma, M. Mälksoo, J. Metsälä, H. Nyysönen és G. Soroka – nézeteinek, természetesen nem felejtkezve el a magyar Gáspár Miklósról, Tóth Juditról, Jakab Andrásról, Halász Ivánról és Schweitzer Gáborról, valamint a már említett Bán Marináról sem. Izsák-Somogyi Katalin az emléktörvényre fogalmat is ad, olvasatában minden törvény és törvényerejű rendelet az, amelynek az Országgyűlés vagy a Népköztársaság Elnöki Tanácsa részéről emlékállítás a célja.⁷ A szerző megállapítja, hogy e tekintetben a tudományos diskurzusban viszonylag nagyobb konszenzus van. Valóban, az általa alkalmazott fogalomhoz az e tanulmányban használt is közel áll, mégis, rá kell mutatnom arra, hogy Izsák-Somogyi Katalin meghatározása túl tág, ami nyilvánvaló lesz alább, Mária királynő 1384. évi dekrétumának ismertetése kapcsán.

Így egy olyan definíciót választottam ki, amit Izsák-Somogyi Katalin nem említ, az ugyanis nem a jogelmélet, hanem a gyakorlati jogélet, a kodifikáció irányából közelít. Az Igazságügyi Minisztérium jogszabályszerkesztéshez készített módszertani útmutatója 1999 júniusában látott napvilágot a *Magyar Közlöny* különszámában. Eszerint az emléktörvény olyan törvény, amely megőrökíti valamely jelentős történelmi esemény vagy kimagasló érdemeket szerzett személyiség emlékét.⁸ A jogalkotó elsődleges célja tehát az, hogy történelmi személyt vagy eseményt *örökítsen meg* az utókor számára. Ebből kifolyólag előfordulhat, hogy a megalkotott törvény kötelező magatartási szabályt, azaz érdemi, valakit valamire kötelező rendelkezést egyáltalán nem tartalmaz,

⁵ Marina Bán: *The Legal Governance of Historical Memory and the Rule of Law*. PhD értekezés. Amsterdam, University of Amsterdam, 2020.

⁶ Izsák-Somogyi Katalin: *A magyar emlék(ezet)törvények rendszere*. PhD értekezés. Budapest, Pázmány Péter Katolikus Egyetem Jog- és Államtudományi Kar Doktori Iskola, 2024.

⁷ Izsák-Somogyi i. m. 48.

⁸ Igazságügyi Minisztérium: *Módszertani útmutató a jogszabályszerkesztéshez*. *Magyar Közlöny*, különszám. (1999. június 28.) 56.

ami elviekben még a jogi jellegét is kétségbe vonja. Ezenkívül lényeges, hogy a megörökített esemény vagy személy *történeti* legyen, vagy közelmúltbeli esemény rögzítéskor annak történetivé válásával már számolni lehessen. Gondoljunk például az 1921 decemberében tartott soproni népszavazásra, amelyet az 1922. évi XXIX. törvénycikk⁹ hagyott emlékül a jövő generációinak. Ám hogy az említett fogalmi elemek valóban együtt legyenek jelen, arra viszonylag későn találunk először példát jogtörténetünkben.

3. Mit tekintünk legelső emléktörvényünknek?

Az *emlékezet* szót már felfedezhetjük középkori joganyagunkban is, elég Mária királynő 1384. évi dekrétumára utalni.¹⁰ Ebben a királynő apjának, Nagy Lajosnak a kiváltságlevelét, amely II. Andrásnak a nemesi szabadság tárgyában kelt oklevelét megerősítve foglalta magában, maga is elfogadta és jóváhagyta. „A minnek örök emlékezetére és ereje céljából” Mária az 1384-es dekrétumát kiadta. Szempontunkból az a lényeges, hogy a dekrétum nem történeti eseményt, hanem *saját*, uralkodói döntést örökít meg, úgy is mondhatnánk, hogy a jogalkotónak a jogszabály tárgyához túlzottan szoros a kötődése. Ilyen alapon ugyanis minden törvény emléktörvény lehetne. Célunkhoz tehát aligha jutunk közelebb, ha az emlékezet vagy hasonló kifejezések használatát kutatjuk régi joganyagok szövegeiben.

Megpróbálhatunk történeti jogtárunkban olyan törvényeket is keresni, amelyeknek semmiféle érdemi rendelkezése sincs. Például az 1548. évi II. törvénycikkben¹¹ „a tőlük kitelhető legmélyebb köszönetet mondják” a királynak „a hozzájuk való véghetetlen jószágáért és kegyességéért” a rendek. Kérték a Mindenhatót rá, „hogy Ő felségét az ő fenséges gyermekeivel együtt napról-napra boldogabbnak, a szerencse minden adományával bőven megáldottnak és ép állapotban levőnek lássák.” Rendkívül érdekes a szintén a protestantizmus terjedése időszakában keletkezett 1553. évi XXIV. törvénycikk,¹² amely meg is indokolja, hogy miért nem „rendelkezik” semmit. Bár „[...] ő felsége – olvashatjuk benne – tévelygő [eretnek] könyvek kiadásáról és nyomtatásáról is megemlékezett, e tárgyban a karok és rendek a maguk részéről azért nem végezhettek semmit, mert nincsen tudásuk róla, hogy fekvő jószágaikban valahol ilyenek lennének.” Pusztán a két példa alapján is világosan látható, hogy az érdemi rendelkezés hiánya nem jelenti azt, hogy valamely törvény bizonyosan az emléktörvények közé sorolható. A kötelező magatartási szabály hiánya ugyanis régen nemcsak az emléktörvények jellemzője volt.

⁹ 1922. évi XXIX. törvény(cikk) a soproni népszavazás emlékének törvénybeiktatásáról. <https://net.jogtar.hu/jogszabaly?docid=92200029.tv>

¹⁰ 1384. évi decretum: II. András és I. Lajos királyok végzeményeinek Mária királynő részéről kettős függő pecsét alatt történt megerősítése. Különösen: 11. §. <https://tinyurl.com/3ee42642>

¹¹ 1548. évi II. törvénycikk: Köszönetnyilvánítás és a király jólétéért való könyörgés. <https://tinyurl.com/2jadfebp>

¹² 1553. évi XXIV. törvénycikk: Az eretnek könyvek kiadása és nyomtatása. <https://tinyurl.com/mrxdaatj>

A XVI. századból mindenképpen meg kell említenünk még az 1567. évi I. törvénycikket,¹³ mivel az igen erős történeti tartalmat hordoz tárgyával összefüggésben. A királynak azon vélekedése nyomán keletkezett, hogy az adott körülmények között célszerűbb fegyverszünetet kötni a törökkel. Az uralkodót bizalmukról biztosították a rendek, hiszen „még élő emlékeztükben tartják [...] a mikor tudniillik Soliman, a törökök fejedelme, az egész Kelet belső erejével együttesen, e legnagyobb és leghatalmasabb hadsereggel, ő felségének eme lesújtott országára tört, hogy ezt végképen elpusztítsa”, a császár és király „az összes külföldi urak és tartományok seregeinek és egész Németországból, a római szent birodalom haderejének idevezénylése mellett személyesen táborba szállott” és magát úgy, mint testvéreit, Ferdinánd és Károly ausztriai főhercegeket „a hadi szerencse minden eshetőségének elszántan alávetette”. A törvénycikk következő szakasza azt is megjegyezte, hogy bár olyan vég helyek, mint Szigetvár és Gyula „a maguk és a nép büneiért a mindenható Isten büntetéséből” elvesztek, annak híre, hogy a császár és király a táborban személyesen van jelen, megakadályozta a törököt abban, hogy az ország megmaradt részét is „tűzzel és vassal végképen tönkre tegye”. Akár már emléktörvénynek is tekinthetnénk e jogszabályt az idézett mondatok miatt, csak hogy azokkal egészen más volt a jogalkotó célja. Nem történelmi eseményt akart rögzíteni, hanem a fegyverszünetet megkötni szándékozó király addigi *érdemeit* kívánta méltatni, amivel már az emléktörvénnyel rokon érdemtörvény típusa körvonalazódik. (Lásd például a Horthy Miklós tíz-, majd húszéves országglását megőrkítő 1930. évi XI. és 1940. évi II. törvénycikkeket.)

Véleményem szerint csak az 1715. évi CXI. törvénycikket¹⁴ lehet az első emléktörvénynek tekinteni. Ebben „örökös köszönetet mondanak” a császárnak és királynak a rendek, hogy az akkori esztergomi érsek kérésére „az érsek úrnak jövendő utódait az esztergomi érseki méltóságban a római szent birodalom igazi hercegeivé” tette. Ezt követően pedig a szövegbe kifejezetten beiktatták, hogy „[...] a tartozó hála jelül e dolog örök emlékezetére azt jelen törvénycikkkel a késő maradékra átbocsájtani akarták...” Vitathatatlan, hogy itt már egy esemény megőrkítése a jogalkotó célja, méghozzá olyané, amelyhez köze aligha van, hiszen a döntést az érsek kérésére az uralkodó és nem a rendi országgyűlés hozta. Ennek megfelelően a törvény érdemi rendelkezést sem tartalmaz. Értelmezésem ellen pusztán annyiban merülhet fel érv, hogy még mindig nincs jelen a maga letisztult formájában az emléktörvény egyik fogalmi ismérveként megjelölt történeti esemény vagy személy. Ám úgy gondolom, hogy a császár és király döntésének történeti jelentőségüvé válásával akár már számolhattak a kortársak, helyes tehát, ha a kezdeteknél kicsit megengedőbbek vagyunk a történetiség vonatkozásában.

A XVIII. század közepétől a XIX. század közepéig a korai emléktörvényeinknek négy típusa jelentkezik. Fent láttunk példát az egyházi vonatkozású döntést megőrkítő

¹³ 1567. évi I. törvénycikk: A király testvéreivel együtt továbbra is arra törekszik, hogy a török császárt az országból vissza nyomja; de a jelen körülmények között az a vélekedése, hogy az ellenséggel fegyverszünetet kell kötni. <https://tinyurl.com/mwkb633k>

¹⁴ 1715. évi CXI. törvénycikk: A karok és rendek Ő legszentségesb felségének köszönetet mondanak, hogy a mostani esztergomi érsek úr hivatalbeli utódait római szent birodalmi herceg méltóságával feldíszíteni méltóztatott. <https://tinyurl.com/ypd6n76n>

törvényre, melyhez hasonlóként érdemes még megemlíteni az 1827. évi XXI. törvénycikket.¹⁵ Ez arról emlékezett meg, hogy Rudnay Sándor hercegprímás munkálkodása eredményeként a primási és érseki széket visszahelyezték Esztergomba, alapításának régi helyére. A további korai emléktörvény típusokat az alábbi címek alatt külön-külön vesszük sorra.

4. Emléktörvényeink alapítványokról

Eger városában a jogászképzés a XVIII. század közepén indult meg, az ottani jogi tanfolyamot „dicséretes és nem csupán hálás emlékezetre, hanem utánzásra is legmél-
több példával” Foglár György, az egri érsek helyettese és főügyésze állította fel. E jeles eseményt a Foglár-féle alapítványról szóló 1741. évi XLIV. törvénycikk¹⁶ rögzíti, azért,

„[...] hogy a fönt említett püspök úrnak ez annyira áldásos szándéka és intéze-
tének [...] a királyi Curiára ruházott legfelsőbb igazgatása állandó emlékezet-
ben maradjon, s ez által is a kellő foganatot annál jobban megnyerje, az ország
karai és rendei jónak látták a mondott alapítvány ügyét az ország törvényeiben
is megemlíteni, s ilyen módon is minden feledségtől (melyet az idők folyása és
viszontagsága magával hozni szokott) megőrizni.”

Bár kétségkívül jogbiztosító értéke is van, olvasatomban e törvénynek mégis inkább felsőoktatásunk történetének egyik kimagasló eseménye megőrkítése a célja. Hiszen a tanfolyam működésének jogi kereteit a Foglár-féle alapítólevél tartalmazza, a törvény erre csupán utal. Az országgyűlésnek ezúttal sincs köze a döntéshez, így – értelemsze-
rűen – az általa alkotott jogszabályban érdemi rendelkezés sincsen.

Hasonló a majdnem száz évvel később keletkezett 1836. évi XXIV. törvénycikk,¹⁷ ami az utókort a Négyesi Akadémiára emlékezteti. Ebben „örök emlékül” törvény-
be iktatták a „jeles és követésre méltó alapítványt” a rendek, mellyel Négyesi Báró Szepessy Ignác pécsi püspök Pécs városában saját költségén bölcselkedést és törvényt tanító intézetet állított fel. Fontos azonban, hogy ilyen típusú jogszabályok nemcsak intézményalapításokhoz kapcsolódtak a korban.

Az 1802. évi IV. törvénycikkben¹⁸ a rendek köszönetüket fejezik ki a királynak, hogy helyesli abbéli szándékukat, hogy amíg sor nem kerül Magyarországon katonai akadémia felállítására, addig a pesti egyetemen szervezzenek egy tanszékot a katonai tudományok előadására. Belefoglalták ezenkívül a szövegbe, hogy tolnai gróf Festetits György negyvezerez, kisérei Rhédey Lajos tízezer, Párnitzky Mihály pedig ezer fo-

¹⁵ 1827. évi XXI. törvénycikk az érseki székeknek s egyszersmind az esztergomi egyház főkáptalanjának alapítása helyére leendő visszaviteléről. <https://tinyurl.com/5n6uafdt>

¹⁶ 1741. évi XLIV. törvénycikk a Foglár-féle alapítványról. <https://tinyurl.com/nn46scrw>

¹⁷ 1836. évi XXIV. törvénycikk Négyesi Báró Szepessy Ignác, Pécsi püspök által Pécsen alapított Akadémiáról. <https://tinyurl.com/4fac3uu3>

¹⁸ 1802. évi IV. törvénycikk a katonai akadémiáról s tanulmányokról, és az ezen czélra gróf Festetits György, Rhédey Lajos és Párnitzky Mihály által tett alapítványról. <https://tinyurl.com/28n5t6t7>

rintot ajándékozott annak érdekében, hogy néhány ifjút – amíg az országban nem állítanak fel hasonló intézetet – hadi tudományokban a bécsi katonai akadémián képezzenek. Megtudjuk továbbá, hogy az uralkodó ígéretet tett rá, hogy az udvari haditanács útján intézkedni fog az új alapítványi növendékeknek a bécsi katonai akadémiára való felvétele iránt. Ebben a jogszabályban találkozunk először a XIX. század első felének nagy nemzeti intézményalapításaihoz kapcsolódó, alább bemutatásra kerülő emléktörvényekben széles körben elterjedt eljárással, hogy a jogalkotó az adományozók nevei mellett az adományozott összegeket is megőrökíti az utókor számára.

1827 – a világhírű zeneszerző, Beethoven halálának éve – jelentős év korai emléktörvényeink történetében: lejjebb részletesen fogok foglalkozni a Magyar Tudományos Akadémia felállításának megőrökítésével. Azonban még az alapítványoknál maradvan megemlítem, hogy az 1827. évi XIX. törvénycikkkel¹⁹ Festetits György fia, László neve is bekerült a jogtárba. Apja említett alapítványát²⁰ ugyanis további tízezer forinttal gyarapította, másrészt szintén tízezer forintot áldozott a Magyar Tudományos Akadémia javára, „névszerint az a mellett a katonai ismerettárgyakra berendezendő mennyiségteni osztály javadalmazására”. A király helyben hagyta a kiadott alapítóleveleket, Festetits László tettét pedig „ezen kiváló bőkezűség és hazafias buzgóság örök emlékezetéül a jelen törvénycikkbe iktatni határozták” a rendek.

Alapítványokról megemlékezhetek utólag is, ennek példája a szintén 1827-ben megjelent XXIV. törvénycikk.²¹ Tomján Margit, Bibics Jakabnak, Arad vármegye első alispánjának özvegye még 1775-ben, Király József pécsi püspök pedig 1814-ben és 1819-ben tett alapítványokat az ifjúság nevelésének és oktatásának elősegítésére. A jogalkotás célját ezúttal is úgy fogalmazza meg a szóban forgó jogszabály, „[...] hogy a hazai ifjúság és tudományok iránt általuk tanúsított bőkezűség emlékezte örökké fönmaradjon: a karok és rendek, Ő szent felsége jóváhagyásával, ez alapítványokat törvénykönyvbe iktatják.” A szövegben Király József neve előtt a néhai szó utal arra, hogy már halott volt a törvény megjelenésének időpontjában. Itt kifejezetten érdekes, ahogy sűrűdik a jogbiztosító érték és az emlékezeti jelleg.

A róla és az alapítványáról szóló jogszabály megjelenését nem érte meg Alagovich Sándor zágrábi püspök sem, aki Pozsega városában alapított árvanevelő intézetet. Csak 1840-ben rögzítette az XLI. törvénycikk,²² hogy „[...] ezen jeles és követésre méltó buzgóságból származott alapítvány örök emlékül törvénybe iktattatik.”

5. A nagy nemzeti intézményalapítások

A fenti példáktól eltérően gróf Széchenyi Ferenc nem pénzbeli alapítványt tett, hanem a magyar nemzet számára könyvtárát, földabroszait – mai kifejezéssel: térképeit –,

¹⁹ 1827. évi XIX. törvénycikk: Gróf Festetits Lászlónak két rendbeli ajánlatai becikkelyeztetnek. <https://tinyurl.com/mr2zw596>

²⁰ A törvénycikk az 1802. évi IV. tc. helyett hibásan a bandériumokról szóló 1808. évi IV. tc.-re hivatkozik.

²¹ 1827. évi XXIV. törvénycikk a Bibits- és Király-féle alapítványokról. <https://tinyurl.com/2w358t2z>

²² 1840. évi XLI. törvénycikk: Néhai Alagovich Sándor zágrábi püspök árvaintézetbeli alapítványa törvénybe iktattatik. <https://tinyurl.com/kyaaamtr6>

illetve ritka pénzérme-, kézirat-, címer- és festménygyűjteményét ajánlotta fel. „Az ország karai és rendei ezen jeles és utánzásra méltó igyekezetnek, hogy emléke az ország köztörvényeiben is fönmaradjon, Ő szent felségének kegyelmes jóváhagyásával becikkelyeztetését elrendelték” – olvashatjuk az 1807. évi XXIV. törvénycikkben,²³ melyben az Országos Széchényi Könyvtár megalapításának az emlékét megörökítették. Jól látható, hogy az 1741-es törvényhez képest a jogbiztosító jelleg „megkopott”, itt az esemény rögzítésén túl már aligha akart mást is a jogalkotó.

Rendkívül fontos a katonai Ludovika-akadémiáról szóló 1808. évi VII. törvénycikk,²⁴ amit ugyan csak részben lehet emléktörvénynek tekinteni, mégis, a benne alkalmazott gyakorlatot a XIX. század első felének számos emléktörvénye követi. E jogszabálynak több érdemi rendelkezése is van, például, hogy valláskülönbőség nélkül, de főképpen nemesi és polgári jogállású személyeket kell felvenni a Ludovikára. A növényedékek száma a százhusz ösztöndíjast és a nyolcvan önköltségest nem haladhatja meg, részben az akadémia pénzalapjára tekintettel, részben, hogy a tanárok sokak miatt ne fordítsanak kevesebb gondot egyesekre. Bár témánk szempontjából mellékes, ez a legelső *numerus clausus* („zárt szám”) jogtörténetünkben, amelyet érdemes összevetnünk a tudományegyetemekre, a műegyetemre, a budapesti egyetemi közgazdaságtudományi karra és a jogakadémiákra való beiratkozást ugyan a felvételt kérők népfiának és nemzetiségének országos arányszámához kötő, de indokolásában a képzés szakszerűségét illetően hasonló elemet is felvonultató 1920. évi XXV., továbbá az azt módosító 1928. évi XIV. törvénycikkkel.²⁵

Főleg tehetősebb „honlakosok” adományaiából gyűlt össze a Ludovika pénzalapja, az adományozókat név szerint és az általuk felajánlott összeggel együtt örökíti meg a törvény rendkívül hosszú 18. §-a. Felsorolását ez a mondat vezeti be: „Hogy pedig azoknak örök emlékezete, a kik ez akadémiára dicséretes buzgósággal ajánlatokat tettek, fönmaradjon: neveik a következő sorban iktattatnak be. Következnek a nevek:” A névsor nem kevesebb mint 240 személyt tartalmaz, a legnagyobb, a kor pénzérték viszonyai szerinti 50.000 forintnyi hozzájárulást Mária Ludovika császárné és királyné (akiről az intézményt elnevezték) adta, őt 30.000 forinttal Albert, Szászország hercege, majd 24.000 forinttal gróf Batthyány Károly királyi kamarás követte a fő támogatók sorában. Évenként fizetendő vagy természetben juttatott felajánlások is előfordultak, például Kamánházy László váci püspök évenként ezer forint fizetését vállalta, továbbá száz akó bort, ötven mérő búzát és ötven öl tűzifát is ajándékozott a Ludovika első beruházására. Bár az évenkénti, tehát jövőbeni juttatások kapcsán még mindig felmerül a jogbiztosító jelleg, a 18. § elsődleges célja – mint fentebb idéztük – nem ez. A névsorban egyébként a természetes személyeken túl a Magyar Hajózási Társaság, a

²³ 1807. évi XXIV. törvénycikk a magyar országos Széchényi-könyvtárról. <https://tinyurl.com/3t7yfbpt>

²⁴ 1808. évi VII. törvénycikk a katonai Ludovika-akadémiáról. <https://tinyurl.com/2vkwyhyv>

²⁵ 1920. évi XXV. törvénycikk a tudományegyetemekre, a műegyetemre, a budapesti egyetemi közgazdaságtudományi karra és a jogakadémiákra való beiratkozás szabályozásáról. <https://tinyurl.com/4xtshf8c>

1928. évi XIV. törvénycikk a tudományegyetemekre, a műegyetemre, a budapesti egyetemi közgazdaságtudományi karra és a jogakadémiákra való beiratkozás szabályozásáról szóló 1920. évi XXV. törvénycikk módosításáról. <https://tinyurl.com/bjddeb8a>

váci káptalan és Lőcse szabad királyi városa is feltűnik, illetve meg lett örökítve egy „névtelen” hármezer forintos adománya is.

A törvény megjelenését követően még többen felajánlást tettek, ehhez bizonyára nem csekély motivációs erőt jelentett, hogy az új adományozók hasonlóképpen számíthattak nevük megörökítésére. A folytatólagos adományozásokról – aminek eshetőségével már számoltak az 1808. évi VII. törvénycikkben – három újabb törvény született, ezek teljes egészükben kizárólag csak neveket és felajánlásokat tartalmazó *emléktörvények*.

Az első, az 1812. évi II. törvénycikk²⁶ az adományozók nevét és az általuk felajánlott összegeket vármegyénként közli, úgy, hogy Horvát- és Szlavónországot, a városokat és a katonai kart (tishti testületeket) is külön-külön csoportokként kezeli. Összesen 137 tételt rögzít, de hogy ebből mennyi a természetes személy, olyan megfogalmazások miatt, mint „Ittebei Kiss Antal, a Kiss Ágost örököseivel együtt” vagy „Baraczházi Kapdebó testvérek” nem lehet pontosan megállapítani. Másrészt a jogalkotó – úgy tűnik – felismerte, hogy csak bizonyos összeghatáron túl van értelme a megörökítésnek, ezért a kisebb felajánlásokat vármegyénként – ahonnet volt vagy ahonnan jelentettek ilyet – összevonta egy tételbe. A legnagyobb összeget, 30.000 forintot ezúttal Pest város közönsége ajánlotta fel. Volt, aki másodsor is adott: például Verhovác Miksa zágrábi püspök, akinek a nevét így törvénybe iktatták újból. Trencsén vármegyéből pedig Sirchich Simon kapitány nemcsak magát, hanem utódait is kötelezte arra, hogy évente kétszáz forintot fizessenek be a Ludovika pénztárába „mindaddig, míg az fönnálland”. Érdekes továbbá a főleg szerbek, illetve horvátok által lakott Szerém vármegyéből érkezett felajánlásra vonatkozó „A kebelbeli nemességtől beszédett és tényleg a pénztárba beszállított 1061 frt” megfogalmazás, ami kérdéseket vethet fel az önkéntesség kapcsán.

A folytatólagos adományozásokat megörökítő második, 1827. évi XVIII. törvénycikk²⁷ – melynek értelmezéséhez a korabeli pénzügyi viszonyokban elengedhetetlenül jártasnak kell lenni –, három névsorral dolgozik. Az első, „A kik a levonást szenvedett bankóczédulákban tett s a fönt idézett törvénycikkekben jegyzett ajánlataiknak váltócédula értékben is, az általuk kijelentett módozat szerint való teljesítésére készeknek nyilatkoztak” – e csoport összesen 47 tételt tartalmaz. Hogy ez pontosan hány személy, azt ugyancsak nehéz összeszámolni egy olyan megfogalmazás miatt, mint „néhai Sternberg Ádám gróf örökösei, nőtestvérei: Sternberg Aloizia s Valburga grófnők, az utóbbi férjezett Salm grófné”. E felsorolás a természetes személyeken kívül említi a magyar királyi kamarai kincstárát és Debrecen szabad királyi városát is. Ezután mindössze két névből áll azok névsora, „A kik vagy az 1811/12-ki törvénycikkek alkotása előtt tettek ajánlatot és még nincsenek becikkelyezve, vagy a kik az 1811/12:II. tc. meghozatala után tettek ajánlatokat a fönt nevezett akadémia javára”. Legvégül pedig azok következnek, akik ajánlataikat váltócédulákban tették meg. E felsorolás összesen 28 tétel, melyek közül kettőben több személy is szerepel, egybe pedig a kisebb, egyen-

²⁶ 1812. évi II. törvénycikk a katonai Ludovika-akadémia részére folytatólag tett ajánlatokról. <https://tinyurl.com/mwny6eyf>

²⁷ 1827. évi XVIII. törvénycikk a Ludovika katonai akadémia javára folytatólag tett ajánlatokról. <https://tinyurl.com/4zsr5mee>

ként száz forinton aluli adományokat foglalták bele név nélkül, összesítve. Csak Thuróczy József ügyvédet emeljük ki e harmadik névsorból, aki arra vállalkozott, hogy azon naptól fogva, ahogy a Ludovikára növendékeket vesznek fel, ezer forint tőke kamatjának címén tizenkét éven át fog fizetni hatvan forintot. Itt is megfigyelhető az, hogy az emléktörvény kötelezettséget rögzít jövőbeni szolgáltatásra.

A folytatólagos adományozásokról szóló harmadik, 1840. évi XLIII. törvénycikk²⁸ négy szakasza négy támogató felajánlásait örökíti meg, újdonsága azonban, hogy a Ludovika részéről immár ellenszolgáltatásokat is megállapít ezek fejében. Az ellenszolgáltatásoknak egyetlen kivétellel az volt a tárgya, hogy milyen hallgatókat kell felvenni az intézményben folyó katonai oktatásra. Nemes Szilágyi Mihály például két alapítványt is tett, de úgy, hogy a Ludovikán abból nevelendő két ifjú egyikének Bihar vármegye, a másiknak pedig Debrecen városa lesz jogosult a kijelölésére. Pest szabad királyi városának támogatását hozzuk másik példának, amely a régi vámház helyét, illetve a közelében lévő városi legelőből egy részt átadott az intézetnek katonai gyakorlóterület céljára, ám azzal, hogy cserében a Ludovika a kertje mögött egy kisebb területet átenged és az új vámház felállítását is állja. A Ludovika-emléktörvények sorozatát ez a jogszabály zárja.

Erősen vitatható, hogy a Nemzeti Múzeumot felállító 1808. évi VIII. törvénycikk²⁹ emléktörvénynek tekinthető-e – álláspontom szerint inkább nem. Talán utolsó mondata indokolja kiváltképp nézetemet: „Egyébiránt a karok és rendek, azon törvényhatóságok emlékezetét, melyek ez intézetet bőkezűséggel elősegítették, föntartani kívánván, határozzák, hogy azok a jövő országgyűlésén az ajánlott összeg megjelölésével, becikkelyeztessenek.” A jogalkotó tehát későbbi időpontra halasztotta az emléktörvény megalkotását. Mégis, az 1808. évi VIII. tc. szövegét olvasva könnyen ellenkező benyomásunk támad. Ugyanis meg lett örökítve benne gyaraki Grassalkovics Antal herceg tette, aki a pesti egyetem fűvészkertje helyett a hatvani kapu melletti saját kertjét átengedte, a cserébe kapott területet pedig a múzeum számára ingyen ajánlotta fel. Ami a törvény legfontosabb érdemi rendelkezését illeti, a nádornak ki kellett jelölnie azt a pénztárat, ahová a múzeumra összegyűlő adományokat be fogják szállítani, a pénzeket elfogadó hivatal pedig köteles volt ezekről éves számadást tenni. Erről sürgősen intézkedni kellett, mert miután a nádor a múzeum alapításának tervét a vármegyéknek megküldte és azokat e célra gyűjtésre szólította fel, számosan „tetemes” ajánlatokat tettek. Olvasatomban tehát e törvényben inkább a szervezés folyamatába tekinthetünk be – nézetem alátámasztására további példaként emlitem, hogy a rendek formálisan felkérték a királyt, hogy szólítsa fel a nádort az építkezés megkezdésére.

²⁸ 1840. évi XLIII. törvénycikk a magyar Ludovica katonai intézetre tett ajánlásról. <https://tinyurl.com/4jm9cctf> A törvénycikk az adakozások törvénybeiktatását illetően az 1808. évi VII. tc. helyett a Nemzeti Múzeum felállításáról, és a magyar nyelv művelését előmozdító más intézkedésekről szóló 1808. évi VIII. tc.-re hivatkozik tévesen.

²⁹ 1808. évi VIII. törvénycikk a Nemzeti Múzeum felállításáról, és a magyar nyelv művelését előmozdító más intézkedésekről. <https://tinyurl.com/mvb965ws>

Az idézett rendelkezés nyomán az 1827. évi XXXV. törvénycikk³⁰ hagyta az utókorra a felajánlók névsorát, ezúttal is minden esetben feljegyezve az adományt. Ez már emléktörvény minden kétséget kizáróan, és igencsak helyeselhető a jogalkotónak az az eljárása, hogy nem az alapító jogszabályt terhelte a felsorolással. A rendek a törvény preambulumban mindenekelőtt József nádornak fejezik ki hálájukat. Az első szakaszban a jogalkotás célját rögzítették: „Hogy pedig a dolognak és azok tanúsított nemeslelkűségének emlékezete fönmaradjon, a kik ez intézetet kész pénzben tett ajánlatokkal, vagy különféle gyűjteményekkel megajándékozták, vagy végül az előbb följánlott összegeket váltóczédula értékre emelték...” A második hosszú, egyben utolsó szakasz előbb 52 vármegyét, majd 40 szabad királyi várost sorol fel – Felső-Bánya kiváltságos várost is beleértve, illetve Selmecet Bélabányával együtt véve –, amit kiegészít további hat tétellel, azaz a Jászok-kunok és a Hajdu városok kerületével, a tizenhat szepesi várossal, Fiume és Buccari tengeri városaival, valamint a kamarai tengeri kerülettel. A pénzadományokat közigazgatási egységenként összesítették és a megfelelő tételhez hozzárendelték. E jogszabályhely egyébként azért is érdekes, mert betekintést ad a korabeli területi közigazgatás szervezetébe. A felsorolásban még 35 tétel következik, melyeknek már jórészt magánszemélyek a szereplői. A legnagyobb, 24.000 forintnyi felajánlást a törvény megjelenése idején már elhunyt gróf Aspermont Linden János Gobertus tette, akinél csak Pest vármegye adott többet. Mint közösségek feltűnnek a „Pesti kebelbeli zsidók”, az „Ó-budai zsidó hitközség”, „Nagy-Körös mezőváros közönsége” és az „Ágostai evangélikus felekezet kis-honti esperessége” is; továbbá voltak magánszemélyek, akik a jövőben évente kívántak meghatározott összeget adni. A megalakuló múzeum gyűjteményét több mint tizenöt fő gyarapította; a pontos számot ezúttal sem lehet meghatározni a „néhai puchói Marczibányi István örökösei” megfogalmazás miatt, és sajnos az sem derül ki, hogy az ajándékozással mely múkin-csek kerültek az intézmény tulajdonába. Ennek a különálló, kizárólag nevekből álló felsorolásnak az élén József nádor áll, a második helyen pedig sárvári felső-vidéki gróf Széchényi Ferenc, egykori királyi főkamrás-mester, aki ekkora elhunyt már.

Folytatólagosan – tehát az 1827. évi törvény megjelenését követően – tett adományok előfordultak a múzeum esetében is, ezek emlékét szintén meg kellett örökíteni. E jogszabályba foglalt kötelezettségre így hivatkozik az 1836. évi XXXVIII. törvénycikk³¹ rövid preambuluma: „A nemzeti Múzeumra 1827-ki Ország-Gyűlés óta tett több-rendbeli nemes lelkű adakozások, az 1808-ik évi 8-dik Czikkely értelmében örök emlékül Törvény-Könyvbe iktattatnak.” Ez az újabb emléktörvény gróf Illésházy István főasztalnokról és felső-eőri Pyrker János László egri érsekről két különálló szakaszt tartalmaz, az előbbi könyvtárát és „diplómái gyűjteményét”, az utóbbi pedig festménygyűjteményét ajánlotta fel a múzeumnak. A harmadik szakasz nyolc újabb pénzadományt jegyez – melyek azonban több mint nyolc személyhez kötődnek –, közöttük gróf Széchényi Lajos tízezer forintos adománya a legjelentősebb, de volt olyan is, aki

³⁰ 1827. évi XXXV. törvénycikk: A magyar nemzeti múzeum javára tett ajánlatok becikkelyeztetnek. <https://tinyurl.com/y4bkf2zs>

³¹ 1836. évi XXXVIII. törvénycikk a nemzeti Múzeumra folytatva tett ajánlásokról. <https://tinyurl.com/yfkuz5fe>

aranyban fizetett. A jogbiztosító érték itt már teljesen eltűnik, a törvény semmilyen jövőbeli kötelezettséget sem rögzít.

Szintén 1827-ben keletkezett az első olyan emléktörvényünk, amely hatályban van ma is: a Magyar Tudományos Akadémia – korabeli kifejezéssel: a tudós társaság – felállítására, illetve a magyar nyelv terjesztésére adományokat tevők neveit örökíti meg az az évi XII. törvénycikk.³² Preambuluma a törvény célját így határozza meg: „Hogy azoknak hálás emlékezete, kik a haza és nyelve iránti kiváló szeretetökből, annak művelésére vagy terjesztésére is nemes szívű ajánlatokat tettek, minden időkre megörökíttessék: a karok és rendek, Ő királyi felségének kegyes jóváhagyása mellett, a törvénykönyvbe iktattnak határozták neveiket...” A következő felsorolás a felajánlók nevét és adományait huszonkét tételben rögzíti, egy helyen azonban az a megfogalmazás, hogy „Széki gróf Teleki József, mostohaanyjának, néhai gróf Teleki László özvegyének, mint az ugyanezen férjtől származott fia természetes és törvényes gyámjának s gondnokának, valamint két testvérének, Teleki Ádám és Sámuel grófoknak is nevökben, legkevesebb 30,000 kötetet magában foglaló könyvtárukat, mely nevezett közös atyjok, gróf Teleki László költségén és fáradozásával gyűjtetett; s azonkívül 5000 forint.” az értelmezést nem segíti. A névsor élén tízezer forintos adományával József nádor áll, nyomban utána következik a legmagasabb, hatvenezer forintos összeggel gróf Széchenyi István. Széchenyi érdemeit természetesen nem csökkentti, de látnunk kell, ahogy cselekedete a nagy nemzeti intézményalapítások – katonai akadémia, nemzeti múzeum – sorába illeszkedik. Továbbá hasonlóan magas, ötvenezer forintos összeget ajánlott fel herceg Batthyány Fülöp is, illetve negyvenezer forintos adományával még gróf Károlyi Györgyöt szintén érdemes fő támogatóként kiemelni. Megemlítjük, hogy idősebb gróf Pálffy Ferenc minden évben négyszáz forintos járadék fizetésére vállalkozott, Szlavniczai Sándor István pedig – aki időközben elhunyt – pénzzel, könyvtárával, éremgyűjteményével, valamint festményekkel és földabroszokkal is megajándékozta a tudós társaságot.

Az 1840. évi XLII. törvénycikknek,³³ melyben „A haza oltárára tett nemeslelkű adakozások örök emlékezetet érdemlő megismerésül, a magyar tudós társaság számára 1832-dik esztendőttől fogva részint kész pénzben, részint kötelezvényben tett ajánlások törvénykönyvbe iktattnak...”, az előbbi törvényhez képest szinte még nagyobb jelentősége van, ugyanis nem kevesebb mint 81 tételt tartalmaz. Megörökített olyan apró adományokat is, mint a „kolosvári unitarium collegium olvasó társasága” által adott öt forint. További érdekesség, hogy szerepel a támogatók sorában egy londoni szücsmester, Mayer János, aki egyébként „szepesi fi”, illetve rögzítésre került két név nélküli, egyenként 88 forintos adomány is. (Bár lehet, hogy e tétel csak hibásan ismétlődik.) Mindazonáltal a három legjelentősebb támogató ezúttal húszezer forinttal gróf Károlyi István, illetve tíz-tízezer forinttal Károly főherceg és gróf Károlyi Lajos volt.

³² 1827. évi XII. törvénycikk: Azoknak nevei, a kik a tudós társaság fölállítására, vagy a hazai nyelv terjesztésére is ajánlatokat tettek, az utókor emlékezete végett törvénybe iktattnak. <https://net.jogtar.hu/jogszabaly?docid=82700012.tv>

³³ 1840. évi XLII. törvénycikk a magyar tudós társaságra folytatva tett adakozásokról. <https://tinyurl.com/5e5juchj>

A felsorolásban vármegyéket és városokat is szép számban találunk, az utóbbiak között például ott van „Sz. kir. Buda fővárosa”, Fiume városa vagy a 16 szepesi város. A Torontál vármegyéből érkezett felajánlásokat a törvény valamilyen okból külön cím alatt közli, a támogatók közül innen példálózva csak a „Török-becsei uradalmat”, a „Nagy-szent-miklósi járást” és a „Nagy-kikindai kerületet” emeljük ki. A csoportosítás sajnos nem következetes: az említett londoni szücsmestert és a kolozsvári olvasótársaságot is e helyen találjuk meg. Ezenkívül figyelmet érdemel, hogy az adományok jellege talán ebben az emléktörvényben a legszínesebb. Míg például báró Szepesy Ignác pécsi püspök évenkénti juttatásként ígért hatszáz forintot, addig Bölöni Farkas Sándor kétszáz darab aranyat adott. Gróf Teleki József kétezer forintot ajánlott fel, ám azt célhoz rendelte: fordítsák a régi pénzgyűjtemény „öregbítésére”. Rendkívül érdekes gróf Pálffy Jánosné, herceg de Ligne Eufemia hagyománya, aki 2500 forintot ígért, melyet csak báró Spiegelné, herceg de Ligne Flóra halálával fizetnek ki a tudós társaságnak. Ez a *mortis causa donatio* római jogi intézményét juttathatná eszünkbe, de e tényállás nem annak felel meg, hiszen itt a halál bekövetkezte, nem pedig a megajándékozott túlélése.³⁴ Ráadásul nem is az ajándékozó, hanem harmadik személy haláláról van szó, akinek bizonyára életében e tőkére és jövedelmére még szüksége volt. Mindez azonban szemléletes példa arra, ahogy az emléktörvényeket is felölölő közjog és a magánjog átszővi egymást a rendi korban. Végül, saját szerzőségű könyveikből adtak hatan, így például Perlaky Dávid *Történeti adatok* című munkájából ötven, Egyed Antal *P. Ovidius Naso pontusi levelei* című fordításából pedig négyszáz példány került az Akadémiára. A Berzsenyi Dániel versei első kiadásából származott tiszta haszonból szintén a hazai tudományosságot képviselő intézmény gyarapodhatott.

6. A „modern” emléktörvény születése

Az utolsó rendi országgyűlés által elfogadott törvények többsége minden jogtörténetben jártas Olvasó számára jól ismert, elég csak megemlíteni olyan törvénycikkeket, mint amelyek a független magyar felelős minisztérium alakításáról (1848. évi III. tc.), az országgyűlési követeknek népképviselőt alapján választásáról (1848. évi V. tc.), a közös teherviselésről (1848. évi VIII. tc.) vagy az ősiség eltörléséről (1848. évi XV. tc.) rendelkeztek. Kevesebben ismerik azonban a legelső jogalkotást, az 1848. évi I. törvénycíkket, amely dicsőült József nádor emlékét iktatta törvénybe.³⁵ Az utolsó rendi országgyűlés ugyanis nemcsak a polgári korszak Magyarországnak alapjait rakta le, hanem – és választott témánk szempontjából ez a fontosabb – új korszakot nyitott emléktörvényeink történetében.

József nádor I. Ferenc osztrák császárnak és magyar királynak volt az öccse, aki az utókortól „a legmagyarabb Habsburg” jelzőt méltán kiérdemelte. Meghatározó szerepet játszott abban, hogy az 1830-as évekre kulturális és gazdasági centrummá fejlődjék Pest városa. Nevéhez fűződik Pest-Buda építészeti arculata, különösen olyan épüle-

³⁴ Földi András – Hamza Gábor: *A római jog története és intéllációi*. Budapest, Nemzeti Tankönyvkiadó, 1996. 650., különösen az 5. jegyzet.

³⁵ 1848. évi I. törvénycikk: Dicsőült József nádor emléke törvénybe iktattatik. <https://tinyurl.com/5ctk7nac>

tek sora, mint a Nemzeti Múzeum, a Gellért-hegyi csillagvizsgáló vagy a Ludovika. Ezenkívül említhetjük még a Városliget rendezését, a Margit-sziget parkosítását vagy a Pest–Vác vasútvonal építését – természetesen mellőzve a teljesség igényét. József nádor adományaival segítette többek között a Nemzeti Múzeum, a Széchényi Könyvtár és a Magyar Tudományos Akadémia megalapítását, illetve számos polgári társaságnak és egyesületnek nyújtott támogatást. Bátyja, a császár és király felé igen hathatósan képviselte a magyar érdekeket: I. Ferencet ő vette rá, hogy az első magyar reformországgyűlést hívja össze, kijárta a kegyelmet a börtönbüntetésre ítélt magyar jakobinusoknak, közöttük Kazinczy Ferencnek, később pedig Kossuth Lajosnak is 1840-ben, valamint a Védegyelet felosztatását szintén ő akadályozta meg. Irányította a mentést az 1838-as pesti árvíz idején, sokat tett a károk enyhítéséért. József nádor Budán és Alsúton élt, díszmagyart hordott és magyarul jól beszélt. Ő volt a Habsburg-Lotaringiai-ház „magyar ágának” megalapítója; kultuszát sajnos megsemmisítették az 1945 utáni időszakban.

Miután József nádor 1847 januárjában elhunyt 71 éves korában, az utolsó rendi országgyűlés elhatározta, hogy a kiváló személyiség emlékét, „aki Habsburgnak születt és magyarnak halt meg” törvénybe iktatja. Mivel csupán három mondatból áll az 1848. évi I. törvénycikk, azt érdemes szó szerint idézni:

„Elhunyt cs. kir. austriai főherceg József, Magyarország nádora, az isteni gondviselés változhatlan végzete által az élők sorából elszolítottván, az ország Rendei nevének a maradék előtt is dicsőített felmutatása végett, Ő Felseége megegyezésének hozzájárultával elhatározták:

1. § József nádor azon ernyedetlen buzgalomért, mellyel félszázadon keresztül, nem ritkán nehéz körülmények között, a hazának ügyeit vezérelte, a nemzet háláját teljes mértékben kiérdemelte.
2. § A történetkönyv azon gyászlapjára, mellyre mulandó életének utolsó órája jegyezve van, az ország Rendei ezennel reá írják nevének a nemzet kebelében el nem múlható emlékezetét.”

A klasszikus emléktörvény jogtörténetünkben ezzel a jogszabállyal született meg: ez az első olyan törvényünk, amely valóban *történelmi személy* emlékét örökíti meg. Semmiféle érdemi rendelkezése sincsen. A dualizmus kori jogtárunkban József nádort már illusztris társaság övezi: Deák Ferenc (1876. évi III. tc.), Andrássy Gyula (1890. évi III. tc.), Erzsébet királyné (1898. évi XXX. tc.) és Ferenc József (1917. évi I. tc.),³⁶ sőt a honalapítás évezredes emlékének törvénybe iktatásával (1896. évi VII. tc.)³⁷ immár megörökítették az első valódi *történelmi eseményt* is. Ezek a jogalkotások azonban már

³⁶ 1876. évi III. törvénycikk Deák Ferenc emlékének törvénybe iktatásáról. [Így!], <https://net.jogtar.hu/jogszabaly?docid=87600003.tv>; 1890. évi III. törvénycikk a gróf Andrássy Gyulának állítandó emlékről, <https://tinyurl.com/2cwk2s2t>; 1898. évi XXX. törvénycikk dicsőült Erzsébet királyné emlékének törvénybe iktatásáról, <https://tinyurl.com/4ehphj2j>; 1917. évi I. törvénycikk Dicsőült Első Ferenc József király emlékének törvénybe iktatásáról, <https://tinyurl.com/yf6sd8ps>

³⁷ 1896. évi VII. törvény(cikk), a honalapítás évezredes emlékének törvénybe iktatásáról. <https://net.jogtar.hu/jogszabaly?docid=89600007.tv>

nem a korai emléktörvényeink körébe tartoznak, hanem jelzik, hogy a XIX. század utolsó harmadában kialakult az emléktörvények kodifikációs gyakorlata.

7. Záró gondolatok

A fenti gondolatmenetet levezethettem volna akár visszafelé is, mégpedig úgy, hogy az MTA emléktörvényből indulok ki. Mert ha ez emléktörvény – amit elismer Tóth Judit is³⁸ –, akkor ezen az alapon a más nagy nemzeti intézményalapításokhoz kapcsolódó hasonló törvényeinket szintén annak kell tekinteni. (Kivétel természetesen van, ilyen például az 1723. évi XLV. törvénycikk,³⁹ amely az országos levéltárat felállította.) Ezek a törvények adományozókról és adományokról emlékeznek meg, szintúgy, mint amelyek egyszemélyi alapítványokat rögzítenek: lényegében tehát csak a bennük szereplő személyek száma tér el. Természetesen nemcsak adományozásokat lehetett megörökíteni, hanem az országgyűléstől független egyéb döntéseket is – az egyházi vonatkozású emléktörvényekhez ezen az útvonalon hasonlóképpen el tudunk jutni. Nézetem szerint ennél korábbi időkben legfeljebb csak emléktörvény-elemeket lehet keresni.

Tóth Judit 2002-ben megjelent programadó munkájának bevezető soraiban olvasható, hogy

„A kormányzat az utóbbi években igencsak magasztos, költséges és nagyszabású rendezvényekkel kényezteti a lakosságot. Pedig az ünnepléseknek hovatovább olyan jelentős igazgatási és jogi rendje teremődik, amely a jogászt is zavarba ejti: vajon a jogról eddig tételezett fogalmat nem kellene-e felülvizsgálnia? [...] A rendi és az ellentmondásokban bővelkedő polgári jogfejlődés egyik rétegét az ünneplés és megemlékezés eszközzé válása jelentette, amely gyakran elszakadt az általánosan kötelező magatartási szabályoktól.”⁴⁰

A szerzőnek a problémafelvetésben vitán felül igaza van: az emléktörvény kérdéssel nyomban fókuszba kerül a jog fogalma. Mindössze arra szeretnék csak rámutatni, hogy ami a kortárs jog tudósát meglepheti, történeti jogtárunkban egyáltalán nem példa nélküli, a jogtörténész számára a kötelező magatartási szabályt nem tartalmazó törvénycikk – némi túlzással – mindennapos eseménynek számít. Mellesleg középkori jogtárunkból néhány ilyen munkájának „A magyarázó törvények” címe alatt maga Tóth Judit is kiemel.⁴¹

³⁸ Tóth Judit: Jelképes jogállam. *Mozgó Világ*, 2002/3. 3–12., különösen: 3.

³⁹ 1723. évi XLV. törvénycikk országos levéltár felállításáról, s mások leveleinek a királyi kamarákból, s a szent Mártonról nevezett szent-pannonhalmi conventból és Erdélyből visszavételéről. <https://tinyurl.com/4f3yfbnf>

⁴⁰ Tóth i. m. 3.

⁴¹ Ezek 1559. évi X. törvénycikk: Reámutatás az okokra, amelyek miatt az országlakók nehezen adóznak. <https://tinyurl.com/238eec73> 1596. évi II. törvénycikk: Az országlakók Ő császári felsége intelméhez képest a haza védelmére ott állának, ha régi erejük és tehetségük megvolna. <https://tinyurl.com/4ez2dem6>

Izsák-Somogyi Katalin 2024-ben megvédett doktori értekezésének mellékletében négy Árpád-kori jogszabályt említ, amit egyértelműen emléktörvénynek minősít.⁴² Véleményem szerint azonban nem sorolható ebbe a kategóriába egyik sem, hiszen egészen más célból keletkeztek. Az általa elsőként citált Szent László király Dekrétomainak Első Könyvének 38. Fejezete ugyanis arról *rendelkezik*, hogy az olyan ünnepeket, mint például karácsony, húsvét, pünkösd vagy éppen Szent Mihály arkangyal napja minden egyháznak meg kell ülni.⁴³ A másodikként hivatkozott 1222. évi I. törvénycikk pedig magát a királyt, II. Andrást *kötelezi* a „szent király ünnepén” fehérvári tartózkodásra, hogy ott még akkor is, „ha mi ott nem lehetnénk, a nádorispán [...] a mi képünkben minden ember dolgát meghallgatja.”⁴⁴ Egészen nyilvánvaló tehát, hogy nem emléket örökít meg e két jogszabály. A kilenc évvel korábbi rendelkezést ismétli meg a szerző által harmadikként jegyzett, székesfehérvári törvénynapokról szóló 1231. évi I.,⁴⁵ illetve az ugyanezt a címet viselő, negyedikként feltüntetett 1267. évi VIII. törvénycikk⁴⁶ – az utóbbiban a bíróság szervezetére („[...] minden megyéből két vagy három nemes tartozik megjelenni...”) érdemes külön figyelmet fordítani. Témánk szempontjából azonban fontosabb a középkori keltezési módra röviden kitérni, mely szerint a dátumot ugyan jelölhették a mai formában (hónap, nap) is, ám ezt lehet a legkevésbé általánosnak tekinteni. Az volt a legelterjedtebb, hogy ha a megjelölni kívánt időpont ünnepnapra esett, akkor dátumként az ünnep nevét tüntették fel, ha pedig arra a napra épp nem esett ünnep, akkor azt jelezték, hogy hány napra történt az esemény a legközelebbi ünnephez képest. Gondoljunk akár az *octavalis* bíróságokra, hiszen azok működése is az ünnepek nyolcadához volt igazítva. A két utóbbi jogszabályban tehát „szent királyunk ünnepe” pusztán dátumírás, szó sincs bennük arról, hogy István király emlékét örök időkre törvénybe iktatnák. A szerző ezután értekezése mellékletében csaknem hatszáz évet ugrik, és az általam vizsgált időszakból – az 1848. évi I. törvénycikket is beleértve – mindössze négy emléktörvényt említ. Pedig mint arra utaltam, a XVI. századból Tóth Judit is hoz joganyagokat, tehát érdemes már kutakodni a török kori jogtárunkban.

Halász Iván és Schweitzer Gábor úgy látják 2018-ban közösen publikált írásukban, hogy

„A kiemelkedő történelmi személyiségek érdemeinek, illetve emlékének törvénybe foglalása egyidős a modern polgári államisággal, illetve a polgári átalakulással. E jogszabályok sorát a József nádor (1776–1847) emlékét megörökítő 1848. évi I. törvénycikk nyitotta meg. [...] Nem tévesztendő össze az érdem-, illetve emléktörvények azokkal a XIX. század első felében elfogadott törvényekkel, amelyek egy-egy nemzeti szempontból kiemelkedő intézmény

⁴² Izsák-Somogyi i. m. 263.

⁴³ Szent László Király Dekrétomainak Első Könyve 38. Fejezet: A szent ünnepek megtüléséről. <https://tinyurl.com/e5n32c9b>

⁴⁴ 1222. évi I. törvénycikk Szent István király ünnepéről. <https://tinyurl.com/2v2ymu4t>

⁴⁵ 1231. évi I. törvénycikk a székesfehérvári törvénynapokról. <https://tinyurl.com/3j6v7nvc>

⁴⁶ 1267. évi VIII. törvénycikk a székesfehérvári törvénynapokról. <https://tinyurl.com/4t8zt8xp>

alapítóinak, vagy támogatóinak felajánlásait sorolták fel, illetve örökítették meg.⁴⁷

A szerzők ezt az álláspontjukat nem indokolják, hanem nyomban megkezdik a József nádorról alkotott emléktörvény tárgyalását. Nézetüket mindenben tiszteletben tartva, és azt is elfogadva, hogy a modern értelemben vett emléktörvényeink sorát az általuk megjelölt jogszabály nyitja, az emléktörvény tágabb fogalma mellett kívánok érvelni jelen munkámmal. Hiszen korai emléktörvényeink is *deklaráltan* emlékezetet örökítenek meg, a kérdés csupán az, hogy ha József nádor történelmi személyiség, akkor Mayer János londoni szücsmester, aki saját anyagi erejéhez mérten az MTA céljára „csak” ötven forintot ajánlott fel, az-e, hiszen az 1840. évi XLII. törvénycikk az ő adományát is „A haza oltárára tett nemeslelkű adakozások örök emlékezetet érdemlő megismeréséül” rögzítette.

Végül, az Igazságügyi Minisztérium által jogszabályszerkesztéshez készített, 1999-ben közzétett módszertani útmutató szerint az emléktörvények megalkotásának kodifikációs módszere Magyarországon a XIX. század végén alakult ki⁴⁸ – megjegyzem, hogy Kampis György kodifikációról írt könyve⁴⁹ ezt az adatot mellözi. A modern emléktörvény kodifikációs gyakorlatának kialakulását magam is a XIX. század utolsó harmadára teszem, amennyiben azonban Olvasóm elfogadja kiterjesztő értelmezésemet, akkor azt is el kell fogadnia, hogy az emléktörvényeknek Magyarországon bizonyos kodifikációs gyakorlata már a XIX. század első felében is létezett.

⁴⁷ Halász Iván – Schweitzer Gábor: Állami és nemzeti jelképek. In: Bába Iván (szerk.): *Közszolgálati protokoll. I. Könyv: Alapvetés*. Budapest, Nemzeti Közszolgálati Egyetem, 2018. 141. A XIX. század első felére vonatkozó bekezdés még nem szerepel a szerzőpáros korábbi, 2010-ben megjelent munkájában, ami említett nézetüknek csak későbbi kialakulására vall. Halász Iván – Schweitzer Gábor: *Szimbolika és közjog. Az állami és nemzeti jelképek helye a magyar alkotmányos rendszerben*. Pozsony, Kalligram, 2010. 143–144.

⁴⁸ Útmutató i. m. 57.

⁴⁹ Kampis György: *Kodifikáció. Elmélet és gyakorlat*. [Budapest], Unió Lap- és Könyvkiadó Kereskedelmi Kft., [1995]. Kampis a jogszabály címe kapcsán csak a Sztálin-emléktörvény hatályon kívül helyezésének esetével foglalkozik: uo. 127.

KÜZDELEM A LETT ÁLLAMNYELVÉRT

Manzinger Krisztián*

1. Bevezetés

A magyar nyelvben az „államnyelv” és az „államnyelvtörvény” kifejezések rossz hangzású, negatív jelentésű szavak. Utóbbira olyan jogszabályként gondolunk, amelynek célja mesterségesen, állami erővel beavatkozni a társadalom nyelvi szokásaiba és valamely domináns, illetve társadalmi vagy regionális szegmensben erős pozíciókkal rendelkező nyelv használatának visszaszorítását elérni az állam hivatalos nyelvének javára. E beidegződés oka az ottani magyar közösség nyelvhasználatát korlátozó és végső célja szerint a nem szlovák nyelvű lakosságot a kétnyelvűség, majd az asszimiláció felé terelni igyekvő – 1995. évi – szlovák államnyelvtörvény, amelyet az elmúlt évtizedekben többször szigorítottak és amely többször volt nemzetközi szervezetek – elsősorban a Velencei Bizottság – vizsgálatának, valamint magyar–szlovák kétoldalú egyeztetéseknek is tárgya.¹ Frissebb, de hasonlóképpen negatív példaként hivatkozható Ukrajna 2017. évi államnyelvtörvénye, illetve általánosságban a 2014

* Egyetemi adjunktus, Károli Gáspár Református Egyetem Állam- és Jogtudományi Kar. ORCID: <https://orcid.org/0000-0001-9244-5741>

¹ A szlovák államnyelvtörvény és általában a szlovákiai nyelvhasználat szabályozása jelentős magyar nyelvű szakirodalmi feldolgozottsággal bír, amelyek közül példaként említhető: Fiala-Butora János: A magyar nyelv jogi helyzete Szlovákiában. In: Kántor Zoltán (szerk.): *Térvesztés és határtalanítás: a magyar nyelvpolitika 21. századi kihívásai*. Budapest, Lucidus-Nemzetpolitikai Kutatóintézet, 2010. 144–171.; Gyuresik Iván: Az államnyelvről szóló törvény a szlovák Alkotmánybíróság döntésének tükrében. *Fundamentum*, 1998/1–2. 40–56.; Horony Ákos: Az anyanyelvhasználat jogi keretei Szlovákiában. *Acta Humana*, 2016/3. 45–55.; Lanstyák István: *A magyar nyelv Szlovákiában*. Budapest–Pozsony, Osiris–Kalligram–MTA Kisebbségkutató Műhely, 2010.; Lanstyák István – Szabó Mihály Gizella (szerk.): *Nyelv. Magyarok Szlovákiában VII. kötet*. Somorja, Fórum Kisebbségkutató Intézet, 2011.; Szabó Mihály Gizella: A szlovák nyelvpolitika és a kisebbségek anyanyelv-használati lehetőségei és törekvései. *Romániai Magyar Jogtudományi Közlöny*, 2006/2. 43–49.; Vogl Márk: A Szlovák Köztársaság államnyelvtörvénye a 2009. évi módosítások után. *Jura*, 2010/1. 126–137.

óta zajló jogkorlátozások és azoknak a kárpátaljai magyarság nyelvhasználati jogaira gyakorolt negatív hatásai Ukrajnán belül és a magyar–ukrán kétoldalú viszonyban.²

Jelen írás a középső balti állam, Lettország itthon kevésbé ismert³ államnyelvi szabályozását kívánja vizsgálni történeti, jogi és társadalmi kontextusban. A vizsgálat különösen indokolt, mert a kárpátaljai magyar közösség nyelvi jogait csökkentő Ukrajna⁴ számára inspiráció és legitimációs forrás is a lett szabályozás: azzal érvelnek, hogy mivel az ukrajnaiakhoz hasonló szabályokat egy már EU-tagállam alkalmaz, így a Kijev által elfogadott jogszabályok nem képezhetnek akadályt Ukrajna EU-csatlakozási folyamatában.

A mostani elemzés ezért először röviden áttekinti az állami nyelvpolitika lettországi történeti alakulását és a legrelevánsabb társadalmi jellemzőket. Ezt követően lett nyelvi szabályozás részleteit, dinamikáját és tartalmát mutatja be a releváns nemzetközi kritikák ismertetésével. Az összegzés nemcsak konkludálja és kontextusba helyezi a lett szabályozást, hanem rávilágít a Lettország és Ukrajna közötti lényeges különbségekre, amely a szerző meglátása szerint kizárja azt, hogy Ukrajna analógiaként tudjon hivatkozni a lett példára saját kisebbségi nyelvi jogokat szűkítő rendelkezései legitímálásának céljával. A vizsgálat a hagyományos jogtudományi módszertant, emellett a történeti módszert, és olykor az összehasonlító módszert alkalmazza, illetve a kutatás elsősorban a szakirodalom feldolgozására és abból következtetések levonására, továbbá a lett jogszabályok és kormányzati dokumentumok elemzésére támaszkodik.

2. Államnyelv az első függetlenség idején

A nyelvi kérdés hagyományosan fontos a lett társadalom számára és jelentős kisebbségi komplexussal párosul annak köszönhetően, hogy a lett nemzet történetének jelentős részében valamely idegen nyelv – évszázadokig a német, illetve a lengyel, majd a 19-20. század fordulóján, illetve a szovjet megszállás időszakában az orosz nyelv – ismerete

² Ezzel kapcsolatosan ld. pl. Ambrus Nikoletta: Az ukrán nyelvpolitika fejlődése és átalakulása 1991-től napjainkig. In: Dabis Attila (szerk.): *Kisebbségi jogok érvényesülése a Kárpát-medencében 2019*. Budapest, Kisebbségi Jogvédő Intézet, 2019. 11–57.; Manzinger Krisztián: Az új Ukrajna és a nemzetiségi kérdés. *Kisebbségvédelem*, 2023/1. 147–150.; Pataki István: A 2012-es ukrán nyelvtörvény és az új nyelvtörvény tervezeteinek összehasonlító elemzése. In: Dabis Attila (szerk.): *Kisebbségi jogok érvényesülése a Kárpát-medencében 2017*. Budapest, Kisebbségi Jogvédő Intézet, 2017. 201–216.; Tóth Mihály: Az anyanyelvhasználathoz való jog az Európai Unión kívül: az ukrajnai példa. *Acta Humana*, 2016/3. 19–31.; Tóth Mihály: Ukrajna Európai Uniói integrációja és a nemzeti kisebbségek jogainak problémája. *Kisebbségvédelem*, 2024/1. 41–98.

³ A szerző tudomása szerint a lett államnyelvi szabályozás rövid bemutatását egy magyarországi szerző végezte el eddig: Bartóki-Gönczy Balázs: Államnyelvek Európában. *Kisebbségkutatás*, 2010/4. 607–609.

⁴ Ennek kapcsán ld. Manzinger Krisztián: Nemzetiségi jogok és stabilitás Ukrajnában. In: Móré Sándor – Szilvássy György Péter (szerk.): *Új védelmi és biztonsági dimenziók az orosz–ukrán konfliktus fényében*. (Jog és állam 49.) Budapest Károli Gáspár Református Egyetem, Állam- és Jogtudományi Kar, 2024. 171–195.

volt a társadalmi felemelkedés feltétele.⁵ Első ízben az ún. Iszkolat-kormányzat – lettországi munkások, katonák és földnélküliek tanácsának végrehajtó bizottsága – tette hivatalos nyelvvé a lettet 1918. január 4-én, az erről szóló rendelet azonban – szükség esetén – engedélyezte idegen nyelvek használatát is a lettel megegyező hivatalos módon.⁶ Alig egy évvel később, a Lett Szocialista Tanácsköztársaság már a több hivatalos nyelv pártján állt, az 1919. március 8-i rendelete az állami és helyi intézmények számára a lett, – az annak történelmi, illetve regionális lett nyelvi változatának tekintett – latgal⁷ és orosz nyelvek használatát engedélyezte egyenrangú módon. Ugyanez a szabályozás a lakosság számára valamennyi „helyi” nyelven, így a fentiekben túl németül, héberül,⁸ litvánul és észtil is biztosította a beadványok megfogalmazásának jogát, kötelezve az intézményeket arra, hogy a beadvány nyelvén válaszoljanak.⁹

A szovjet hatalom elsöprését követően az 1920-as évek elején konszolidálódott Lettországnak kezdetben nem volt nyelvtörvénye, az 1934-es puccsig a lett mellett a német és az orosz is hivatalosnak számított az országban. Ezeknek az ismeretét megkívánták a lakossággal érintkező hivatalnokoktól, használható volt a parlamentben, sőt, egy 1921-es rendelet hivatalos státuszt biztosított a latgalliaknak is.¹⁰ Az ország etnikai összetétele a maitól eltérően nézett ki: az 1920-as első népszámlálás szerint a lakosság 72,7%-a volt lett, 7,8%-a orosz, 4,9%-a zsidó, 4,7%-a belarusz, 3,6%-a német, 3,4%-a lengyel, 1,6%-a pedig litván nemzetiségű.¹¹ A városi lakosság jelentős része német nyelvű volt, míg a lettek és oroszok elsősorban vidéki lakosok voltak.¹² Az 1930-as évektől egyre nagyobb figyelem irányult a hagyományosan rengeteg dialektusban beszélt lett nyelvre, és akkor fogalmazódott meg erős politikai igény annak „megtisz-

⁵ Sergei Kruk: Evicting the Speaking Subject: A Critique of Latvian Concepts of Language. *Journal of Baltic Studies*, 2011/4. 447.

⁶ Iskolata dekrēts par latviešu valodas lietošanu Latvijas iestādēs. 1918. január 4., <http://www.vvk.lv/index.php?sadala=135&id=170>

⁷ A balti nyelvcsaládhoz tartozó, a lettől éles határokkal el nem választható dialektus, amelyet egyesek önálló nyelvnek, mások a lett regionális nyelvjárásának tekintenek. Kialakulásának legfőbb oka történelmi, amíg Lettország többi részétől eltérően a latgallnak helyt adó térség nem német, hanem lengyel fennhatóság alatt állt, így a német nyelvnek kevésbé, a szomszédos keleti szláv dialektusok hatásainak viszont annál inkább kitett volt. Văč Ivanov (Zoltán András fordítása): A szláv nyelvjárások és a Litván Nagyfejedelemség egyéb nyelvei, *Kisebbségkutatás*, 2003/4. 818.; Tomas Kamusella: *Words in Space and Time – Historical Atlas of Language Politics in Modern Central Europe*. Budapest–Vienna–New York, Central University Press, 2022. 111. Latgallia kapcsán ld. Manzinger Krisztián: Lettország megtalált harmadik csillaga: Latgallia múltja és jelene. *Pro Minoritate*, 2024/3. [a továbbiakban: Manzinger (2024a)] 14–31.

⁸ A lettországi zsidó közösség tudatosan használta a hébert az oktatásban és a közéletben is, míg a jiddisnek csak marginális, kiegészítő szerepet adtak. *Uj Kelet*, 1925/248., 1925. október 30., 8.

⁹ Latvijas padomju valdības dekrēts par oficiālos rakstos lietojamām valodām. 1919. március 8. <http://www.vvk.lv/index.php?sadala=135&id=169>

¹⁰ Noteikumi par latgaliešu izloksnes lietošanu. 1921. augusztus 17. <http://www.vvk.lv/index.php?sadala=26&id=238>

¹¹ Ethnicities in region of Latvia. Statistics. <https://tinyurl.com/4949asss>

¹² Észtország, Lettország és Litvánia német és orosz kisebbségek. *Stud Statistiskai Tudósító*, Vol. 8., N. 219. (1939. október 16.) 3.

títása” és standardizálása iránt.¹³ Ez – hasonlóan a két másik balti államhoz – jelentős nyelvújítást és nyelvfejlesztést jelentett, amelynek egyik kiemelt célja az orosz nyelv-tani hatások eltávolítása volt.¹⁴

Végül a Kärilis Ulmanis vezette 1934. május 15-i puccsot követően, 1935 januárjában született meg az első lett államnyelvtörvény.¹⁵ Ez a jogszabály a lett használatát kötelezően előírta a hadseregben, az állami és helyi intézményekben, valamint a vállalatok számára, illetve közjogi jelleggel bíró magánintézmények számára. A törvény megkövetelte a latin betűk használatát is – ebben érdemes látni a cirill és különösen a gót betűk¹⁶ elleni küzdelmet. Az egyetlen hivatalos nyelvként való rögzítés mellett a törvény 4. cikke lehetővé tette az egyén és a hivatal kapcsolattartásában a német és az orosz nyelv használatát azokon a településeken – minden település esetében külön belügyminiszteri engedélyhez kötve –, ahol a legutóbbi népszámlálás szerint az adott nemzetiség aránya meghaladta az 50%-ot. A törvény szabad nyelvhasználatot biztosított – amennyiben az nem sértett lett jogszabályt – a nyilvánosság számára zárt üléseken, a kereskedelemben, az egyházakon belül, a sajtóban és az agráriumban.

A lett elsőségét jól mutatták az olyan előírások, miszerint a helyi önkormányzatokban németül vagy oroszul elhangzó hozzászólásokat már egy fő azt igénylő esetén lettre kellett fordítani, illetve, hogy a nyilvánosságnak szóló bármilyen tájékoztatást kötelezően lettül kellett kiírni, míg más nyelv használatát belügyminiszteri engedély esetén tették lehetővé. A szándékos jogsértéshez szankciót is rendelt a jogalkotó: 1000 lat¹⁷ pénzbírságban vagy 3 hónap elzárásban maximalizálva azt, amelyeket együtt is ki lehetett szabni. Az állami fellépés azonban nem volt drasztikus: az orosz kisebbség nyelvi jogai nem sérültek lényegében,¹⁸ habár a társadalmi integráció legfontosabb eszközének innentől a jó lett nyelvtudást tekintették.¹⁹

3. A szovjet megszállás hatása a nyelvi tájképre

A Szovjetunióban az 1930-as évek második fele óta egyre erősebb volt az etnikumközi kommunikáció nyelvének nevezett orosz dominanciája, ami jelentős mértékben gyengítette minden más nyelv pozícióját, így már nagyon rövid távon a lettét is. Miközben az 1944 őszen visszatérő szovjet hatalom sürgette a nagy számban betelepülő orosz,

¹³ Kruk (2011) i. m. 448–449.

¹⁴ Ina Druviete: Linguistic human rights in the Baltic States. *International Journal of the Sociology of Language*, Vol. 127. (1997) 164–165.

¹⁵ Likums par valsts valodu. 1935. január 9. <http://www.vvk.lv/index.php?sadala=135&id=166>

¹⁶ Lettország társadalmi elitjének jelentős részét ekkor még a balti németek tették ki, a korszakban pedig elterjedt volt a gót betűk használata német nyelvterületen.

¹⁷ Ez jelentős összegnek, mintegy 40 akkori brit fontnak számított, a korszak legnagyobb címletű lett bankjegye az 500 latos volt.

¹⁸ Irina Saburova: The Soviet Occupation of the Baltic States. *The Russian Review*, Vol. 1. (1955) 36–37.

¹⁹ Liliija Alijeva: Left Behind? A Critical Study of the Russian-speaking Minority Rights to Citizenship and Language in the Post-Soviet Baltic States. Lessons from Nationalising Language Policies. *International Journal on Minority and Group Rights*, Vol. 4. (2017) 495. <https://doi.org/10.1163/15718115-02404004>

hogyan tanuljanak meg lettül, ismerjék meg az ország történelmét és kultúráját, és ennek érdekében még tananyagok elkészítését is megrendelte,²⁰ a betelepülők és a hivatalos szervek működése miatt az orosz gyorsan nyert teret, és az 1950-es évekre már az is gyakorlattá vált, hogy a lett tanórákon oroszból fordított szövegeket, orosz irodalmat és történelmet tanultak a diákok.²¹

A lett nyelv védelmében a két világháború közötti időszakot jelentős részben Moszkvában, illetve Moszkva megbízásából illegálisan Lettorszáiban töltő Jānis Kalnberziņš vezette Lett Kommunista Párt megpróbált fellépni: 1956-ban olyan rendelkezést hoztak, amely a munkahelyük elvesztésével fenyegette azokat, akik két éven belül nem tanulnak meg lettül. Nem sokkal később, 1958 januárjában a XV. lett pártkongresszuson nyilvános felszólalások sürgették az orosz mellett a lett nyelv használatára vonatkozó szabályok betartását is, arra hivatkozva, hogy anélkül tömeges politikai munka nem végezhető.²² Végül a hrucsovi pártvezetés az 1959-es szovjet oktatási reformmal való szembe szegülés miatt – amelynek egyik újítása a lett nyelv Lett Szovjet Szocialista Köztársaság (SZSZK) valamennyi iskolájában való kötelező tanításának „önkéntes” eltörlése lett volna²³ – eltávolította a helyi pártvezetést;²⁴ Kalnberziņš utódja a két világháború közötti időszakot moszkvai pártmunkával töltő *homo sovieticus* Arvīds Pelše lett.

A következő évtizedek a nyílt oroszosításról szóltak: az 1970-es évekre a Lett SZSZK-ban volt – az Oroszországi Szovjet Szövetségi Szocialista Köztársaságot (OSZSZSZK) leszámítva – a legmagasabb valamennyi tagköztársaságban az iskolákban orosz nyelven tanított órák száma,²⁵ sőt a Lett SZSZK-ban lett tannyelvű iskolába járók az orosz iskolákba járóknál egy évvel tovább, 10 helyett 11 évig voltak diákok, hogy megfelelően megtanulják az oroszot.²⁶ Erre az időszakra – más tagköztársaságokhoz hasonlóan – a Lett Tudományos Akadémián belül létrejött olyan részleg, amelynek az orosz nyelv oktatásának előmozdítása volt a feladata,²⁷ míg az 1980-as évek pártiratai ezen felül is az orosz nyelv oktatásának további erősítését szorgalmazták a nemzetek közötti kommunikáció és az internacionalizmus előmozdítása érdekében.²⁸

²⁰ Jānis Riekstiņš: *In Defence of the Latvian Language Against Russification – Documents from 1944-1989*. Riga, Latviešu valodas aģentūra/The Latvian Language Agency, 2012. 22–23.

²¹ Yaroslav Bilinsky: Education of the Non-Russian Peoples in the Soviet Union. *Comparative Education Review*, Vol. 1. (1964) 82.

²² Jacob Ornstein: Soviet Language Policy: Theory and Practice. *The Slavic and East European Journal*, Vol. 1. (1959) 16–17.

²³ Bilinsky i. m. 84.

²⁴ A folyamat okairól, lefolyásáról és következményeiről ld. Michael Loader: A Stalinist Purge in the Khrushchev Era? The Latvian Communist Party Purge, 1959–1963. *The Slavonic and East European Review*, Vol. 2. (2018) 244–282. <http://dx.doi.org/10.5699/slaveasteurorev2.96.2.0244> A folyamat levezénylője Arvīds Pelše volt.

²⁵ Roman Solchanyk: Russian Language and Soviet Politics. *Soviet Studies*, Vol. 1. (1982) 31.

²⁶ Gabrielle Hogan-Brun: Language-In-Education across the Baltic: Policies, Practices and Challenges. In. *Comparative Education*, Vol. 4. (2007) 561.

²⁷ Solchanyk i. m. 29.

²⁸ Riekstiņš i. m. 80–89.

A kiemelt figyelem oka azonban nem az volt az, hogy a lett társadalmat ne sikerült volna a kétnyelvűség irányába eltolni: 1970-ben a lett nemzetiségűek 4,5%-a, 1979-ben már 4,8%-a beszélt az orosz anyanyelvként, míg 45,1%-a, valamint 56,7%-a folyékonyan, második nyelvként – ez utóbbi adat Belorusszia után a legmagasabb arány volt valamennyi tagköztársaság között – nem számítva az OSZSZSZK-t.²⁹ Végül az 1989-es utolsó szovjet népszámlálás adatai szerint a lett nemzetiségű lakosság 69%-ának volt anyanyelve vagy folyékonyan beszélt második nyelve az orosz, ami alig maradt el az Ukrán SZSZK-ban élő ukránok 73%-ától, és jóval meghaladta az észtek 35%-os és a litvánok 38%-os arányát.³⁰ Ugyanekkor a teljes lettországi lakosság 37,7%-a nem beszélt a lettet – ez a tituláris nyelvek vonatkozásában Kazahsztán és Kirgizisztán után a harmadik legmagasabb arány volt a Szovjetunióban³¹ –, döntően amiatt, hogy a népesség harmadát kitevő oroszoknak csak 22%-a beszélt lettül. Az orosz a más lettországi kisebbségek között is dominánssá vált: a fehéroroszkok 2,5%-a beszélt lettül és 94,5%-a oroszul, míg az ukránok esetében 0,9%, illetve 93,2%, a lengyelek esetében 27,5%, illetve 87,9%, a litvánok esetében 23,8%, illetve 47,9% volt az aránypar 1989-ben.³²

4. A lett államnyelvi szabályozás a függetlenség visszaállításának idején

Mivel a nemzeti nyelv visszaszorulásának folyamata a szovjet megszállás évtizedeiben kollektív társadalmi traumát okozott,³³ a Lett SZSZK volt az első, ahol 1986 októberében, a *glasznosztj* idején felvetődött az oktatási reform kérdése, természetesen a lett nyelv pozíciójának erősítése céljával. Az egyidejűleg megindult politikai szerveződés etnikai polarizációt eredményezett: az egyes tagköztársaságokban létrejött népfrontok saját nemzeti közösségük védelmének szándékával etnikai alapon teremtettek kihívást a kommunista pártnak, erre pedig a tagköztársaságokban a hatalmukat féltők, első sorban az oroszok, szintén etnikai szerveződéssel reagáltak.³⁴ A nyelvi vita különösen kiéleződött Lettországbán: egyes lettországi oroszok 1989 februárjában Mihail Gorbacsov szovjet pártfőtitkárnál tiltakoztak a lett nyelv hivatalos nyelvvé tételével szemben, azzal érvelve, hogy a tagköztársaság lakosságának 50%-a nem lett.³⁵ Tiltakozásuk nem

²⁹ Michael Kirkwood: 'Glasnost', 'the National Question' and Soviet Language Policy. *Soviet Studies*, Vol. 1. (1991) 68–69.

³⁰ David F. Marshall: The role of language in the dissolution of the Soviet Union. *Work Papers of the Summer Institute of Linguistics, University of North Dakota Session*, Article 3, 1992. 45.

³¹ Paul Kolstoe: *Russians in the Former Soviet Republics*. Bloomington and Indianapolis, Indiana University Press, 1995. 89.; Triin Vihalemm – Gabrielle Hogan-Brun: Language policies and practices across the Baltic: processes, challenges and prospects. *European Journal of Applied Linguistics (EuJAL)*, Vol. 1. (2014) 59. <http://dx.doi.org/10.1515/eujal-2013-0004>

³² Ilmárs Mežs – Edmunds Bunkše – Kaspars Rasa: The Ethno-Demographic Status of the Baltic States. *GeoJournal*, Vol. 1. (1994) 18.

³³ Druviete (1997) i. m. 167.

³⁴ Marshall i. m. 41.

³⁵ Riekstiņš i. m. 145–148.

járt sikerrel, de a balti államok közül Lettország volt az utolsó, amely 1989 májusában – az orosz mellett – párhuzamosan hivatalosnak nyilvánította a nemzeti nyelvét.³⁶

Az 1989-ben elfogadott nyelvtörvény³⁷ az alkalmazott kétnyelvűséget ültette volna gyakorlatba,³⁸ viszont eleve úgy rendelkezett, hogy a legfontosabb előírásai csak három évvel később, 1992 májusában lépnek hatályba. Az 1989-es szöveg az állampolgárra bízta annak eldöntését, hogy a hatóságokhoz lett vagy orosz nyelven fordul, így az intézmények és azok alkalmazottai kötelezetté váltak mindkét nyelv megfelelő ismeretére. A hely-, intézmény- és szervezetnevek hivatalos változatának a lettet tette, a más nyelvű átiratok elkészítését a helyi tanácsokhoz utalta. Az egyes pozíciók betöltéséhez szükséges lett nyelvtudás méréséhez egy háromfokozatú skálát vezetett be, a hivatalnokok számára pedig hároméves átmeneti időszakot és munkaidőben szervezett nyelvtanfolyamokat biztosított a lett nyelv elsajátítására.³⁹

1992. március 31-én azonban, még a hatálybalépés előtt, de már a függetlenség visszaállítása⁴⁰ után a jogalkotó olyan mértékben megváltoztatta a szöveget, hogy egyes vélemények szerint az már új törvénynek volt tekinthető.⁴¹ A módosítás deklarált céljai az előrehaladott nyelvcseré folyamatának megállítása, az állam kétnyelvűvé válásának megakadályozása és a közéletbeli nyelvi hierarchia megváltoztatása voltak. A szabályozás részletei tekintetében a quebeci és a katalóniai mintákból merítettek inspirációt a lett jogalkotók.⁴² Az új jogszabályszerző megszüntette az állam kétnyelvűségét, a preambulumban deklarált céllá pedig a megelőző évtizedek beszűkült nyelvhasználati lehetőségei okán az államnyelvvé nyilvánított lett pozícióinak megerősítése vált a köz- és állami élet, illetve az oktatás területén.⁴³ A preambulumban azt is rögzítette, hogy Lettország tisztelettel kezeli az országban használt valamennyi nyelvet és dialektust, illetve, hogy az államnyelvvé nyilvánítás nem érinti a más nemzetiségű lakosok alkotmányos jogát anyanyelvük vagy más nyelv használatára.

A törvény 4. cikke már csak a lett nyelv kötelező ismeretét írta elő az ügyfelekkel érintkező állami alkalmazottaknak – az 1989-es jogszabály hatálybalépésének három évvel való halasztását utólag pont ezzel indokolták, vagyis, hogy a lettül nem tudó állami alkalmazottaknak legyen elég idejük megtanulni a nyelvet, amit az állam ingyenes, munkaidőben zajló nyelvtanfolyamokkal segített elő⁴⁴ –; más nyelvek vonatkozásában

³⁶ Kirkwood i. m. 75.

³⁷ A jogszabály eredeti, 1989-ben elfogadott változatát angolul közli Riekstiņš i. m. 149–153.

³⁸ Jacques Maurais: Regional majority languages, language planning, and linguistic rights. *International Journal of the Sociology of Language*, Vol. 127. (1997) 153.

³⁹ Druviete (1997) i. m. 127., 167.

⁴⁰ Lett alkotmányjogi érvelés szerint az ország függetlenségének 1940-es szovjet megszállás következtében történt elvesztése jogellenes volt, ezért a Szovjetunió széthullásának folyamatában az állam ezt a jogellenes állapotot szüntette meg, amikor jogfolytonosan visszaállította a függetlenségét.

⁴¹ Kolstoe i. m. 126.

⁴² Ina Druviete: Language Policy and Protection of the State Language in Latvia. 2002. 1–2. <https://tinyurl.com/4emwam6c>

⁴³ A törvény szövege elérhető Par grozījumiem un papildinājumiem Latvijas Padomju Sociālistiskās Republikas Valodu likumā. <http://www.vvk.lv/index.php?sadala=135&id=165>

⁴⁴ Druviete (2002) i. m. 2.

a munkájuk ellátásához szükséges mértéket tette mércének. A 6. cikk az állami intézményekben belső nyelvként az államnyelvet rendelte használni, ha valaki azt nem beszélte, akkor megegyezés alapján nyílt lehetősége más nyelv használatára, de ilyenkor mindent lettre kellett fordítani, ha csak egy jelenlévő is kérte. Állami és helyi önkormányzatok ügyintézési nyelve csak a lett lehetett és a köztük levő érintkezés is csak államnyelven folyhatott. Az állam által kibocsátott dokumentumok lett nyelven készültek – az oktatásban az angol második nyelvként használatát lehetővé tette a jogalkotó –, míg az állami intézmények kötelesek voltak a lett mellett angol, német vagy orosz nyelven is beadványokat elfogadni, amire kötelesek voltak lettül válaszolni, de mellette adhattak választ a kérvény nyelvén is (8. és 9. cikkek).

Kötelezővé tették a nem lett nyelvű rádió- és televízióadások lett tolmácsolását, illetve feliratozását. A helyneveket lettül kellett feltüntetni, az eltérő nyelvű kiírások lehetőségét a jogalkotó a helyi önkormányzatok hatáskörébe utalta (16. és 17. cikkek), míg az állami intézmények, önkormányzatok és jogi személyek bélyegzői kizárólag lett nyelvűek lehettek. A hirdetések kötelezően lett nyelven készülhettek, más nyelvet a lett nyelv után lehetett használni. A 15. cikk elismerte és védelemben részesítette az írott latgal nyelvet, illetve a lív nyelvet.

A törvénytől a jogalkotó a lettországi társadalom „újraegyesítését” várta annak révén, hogy a lett nyelv kötelező használatának előírása nyelvi, majd etnikai hasonulást, vagyis homogenizálódást eredményez.⁴⁵ Ez irreális elvárás volt, azonban az állam inentől kezdve igyekezett ennek az intézményi kereteit megteremteni, különösen az 1992-ben létrehozott, a törvény végrehajtását felügyelő Államnyelvi Központ (VVC) által. Ugyanakkor bevezették a törvény megszegésének szankcionálására szolgáló bíróságok intézményét is – ezt az EBESZ akkori kisebbségi főbiztosa, Max van der Stoel kifejezetten kritizálta a rigai kormánynak küldött 1993. április 6-i levelében.⁴⁶

5. A hatályos szabályozás, az 1999. évi lett államnyelvtörvény

Az 1992-es államnyelvtörvény után szektorális jogszabályokban is erős hangsúlyt kapott a lett nyelv pozícióinak erősítése. Különösen igaz volt ez az állampolgárság-⁴⁷ és oktatáspolitikai területén,⁴⁸ illetve a médiaszabályozás vonatkozásában: az 1995-ös médiatörvény 1997-es és 1998-as módosításait követően a Nemzeti Rádió és Televíziótanács sok rádió és tv-csatorna működését szüntette be, amelyek a törvényileg

⁴⁵ Sergei Kruk: Latvian Concept of Linguistic Integration. In: Stephan Kessler (szerk.): *Contributions to Baltic-Slavonic Relations in Literature and Languages*. Berlin, Logos Verlag, 2022. 74.

⁴⁶ Kolstoe i. m. 126.

⁴⁷ Manzinger Krisztián: Az állampolgárság mint nemzetépítési eszköz – Lettország a jog, a politika és történelem háromszögében. *Állam- és Jogtudomány*, 2024/3. [a továbbiakban Manzinger (2024b)] 46–50. <https://doi.org/10.51783/ajt.2024.3.03>

⁴⁸ Manzinger Krisztián: Asszimiláció vagy integráció – A lett oktatáspolitikai célrendszere. *Kisebbségi Szemle*, 2024/1. [a továbbiakban Manzinger (2024c)] 71–86. <https://doi.org/10.62153/ksz.2024.1.mk>

előírt 20%-os maximumnál nagyobb arányban sugároztak nem lett nyelven – elsősorban oroszul.⁴⁹

Az 1999-es új államnyelvtörvény⁵⁰ kontextusát elsősorban mégis a lett államnyelvi státuszának 1998. november 6-i alkotmányos szintre emelése és az oktatási törvény módosítása jelentette, amely a kétnyelvű tanítás bevezetésével a kisebbségi tannyelvű iskolákban 2004-re megszüntetni rendelte a szovjet időkből megörökölt párhuzamos lett és orosz nyelvű oktatási rendszereket. Mindezen változtatások oka az volt, hogy habár a lett nyelv társadalmi pozíciói erősödtek az 1990-es években, így a 2000-es népszámlálás szerint az ország lakosságának már 82%-a beszélt lettül, szemben az 1989-es 62%-kal,⁵¹ a kormányzat mégis elégedetlen volt B1-es szint elérését célzó az állami nyelvtanulási programokkal, amelyekbe 1996-tól az ENSZ Fejlesztési Programja (UNDP) is bekapcsolódott.⁵²

A 2000. szeptember 1. óta hatályos államnyelvtörvény preambuluma a jogszabály céljait több vonatkozásában definiálja: szolgálja a lett nyelv védelmét és fejlesztését, a lett nép kulturális és történelmi örökségének megőrzését, a lett nyelv szabad használatának biztosítását bármely kontextusban az ország területén, a kisebbségek nyelvhasználatára szabadságának biztosítása mellett a lett nyelv ismerete és használata révén társadalmi integrációjuk előmozdítását, a társadalmi kohézió gyorsítását, illetve a lett nyelv szerepének a kultúrában való erősítését. A törvény tárgyi hatályát a 2. cikk tartalmazza, amit a könnyebb érthetőség kedvéért érdemes visszafelé haladva kibontani. A (3) bekezdés szerint nem esik a törvény hatálya alá a lettországi lakosok informális beszélgetése, a nemzeti és etnikai csoportok belső kommunikációja, illetve az egyházi nyelvhasználat. A (2) bekezdés szerint magánintézmények, szervezetek, cégek és önfoglalkoztatók tevékenysége akkor esik a törvény hatálya alá, ha közérdeket érint – például köz- és munkabiztonság, egészségügy, illetve fogyasztói és munkavállalói jogok területét –, és akkor is csak a szükséges mértékben. Az (1) bekezdés alapján kifejezetten a törvény hatálya alá tartozik a központi és helyi intézmények, bíróságok és más közintézmények, továbbá igazságszolgáltatási és oktatási intézmények nyelvhasználatára.

A 3. cikk (1) bekezdése a lettet államnyelvként deklarálja, majd a (2) bekezdés biztosítja a jogot, hogy az ország területén bárki bármilyen intézményhez az államnyelven forduljon. A 3. cikk (4) bekezdés a latgal írott változatának, a 4. cikk pedig a lívnek az állami védelmét deklarálja. Az 5. cikk – összeolvasva a 3. cikk (4) bekezdéssel – a lív kivételével minden más nyelvet idegennek nyilvánít az ország területén.⁵³ A 23. cikk (1) szerint a hivatalos kommunikációban csak a lett hivatalos terminológiáját lehet

⁴⁹ A vonatkozó rendelkezéseket végül az Alkotmánybíróság 2003. június 5-i döntésében alkotmányellenesként megsemmisítette, amit az ET keretegyezményének tanácsadói bizottsága is üdvözölt. Advisory Committee on the Framework Convention for the Protection of National Minorities, Opinion on Latvia adopted on 9 October 2008. ACFC/OP/I(2008)002, 94. pont, 25. o. <https://tinyurl.com/4nc2xfkd>

⁵⁰ A törvény szövege lettül elérhető: Valsts valodas likums, <https://tinyurl.com/ye23z577>

⁵¹ Hogan–Brun i. m. 557.

⁵² Kristine Krūma: *Country Report on Citizenship Law: Latvia*. EUDO Citizenship Observatory, 2015. 18.

⁵³ A latgal nem önálló nyelv lett értékelés szerint, hanem a lettnek egy regionális nyelvváltozata.

használni – vagyis a *latgalt* nem.⁵⁴ A 6. cikk (1) bekezdés az állami és helyi intézmények, bíróságok, igazságszolgáltatási intézmények, továbbá többségi állami és önkormányzati tulajdonú vállalatok munkavállalóinak kötelezővé teszi a munkavégzésükhöz szükséges mértékben a lett nyelv ismeretét és használatát. A (2) bekezdés szerint a magánintézmények, a (3) bekezdés szerint a törvény alapján közfeladatot ellátó magánintézmények, a (4) bekezdés szerint a külföldi vállalatok Lettorszáiban dolgozó szakemberei a munkájuk végzéséhez szükséges mértékben szintén kötelesek a lett nyelvet ismerni és használni. Habár a munkavállalókra vonatkozó szabályozást és a hozzá kapcsolódó gyakorlatot a keretegyezmény végrehajtását monitorizáló Tanácsadó Bizottság folyamatosan kritizálja, a lett jogalkotó az elmúlt évtizedekben inkább kiterjesztette a kötelezettséget és mindössze a vasúti dolgozók esetében történt némi enyhítés a 2020-as évek elején az aggasztó munkaerőhiány csökkentése okán.⁵⁵

A 7. cikk (1) és a 8. cikk (1) bekezdései értelmében az állami és helyi intézmények, bíróságok, igazságszolgáltatási intézmények, állami vagy önkormányzati többségi tulajdonú cégek belső ügyviteli nyelve a lett, s amennyiben szükséges szóbeli belső kommunikációban valamilyen idegen nyelv használata, akkor gondoskodni kell a lett nyelvű tolmácsolásról, míg külföldi országok intézményeivel való érintkezésnél használható idegen nyelv. A 10. cikk értelmében minden intézménynek, szervezetnek és vállalatnak kötelessége lett nyelvű beadványt elfogadni, sőt, idegen nyelvű dokumentumot csak akkor fogadhat el, ha ahhoz mellékeltek államnyelvi fordítást, vagy ha külföldről érkezik – ebben az esetben nem kell fordítást csatolni hozzá. A kivételeket taxatív jelleggel a 10. cikk (2) bekezdés tartalmazza, amely rendőrségi és egészségügyi bejelentéseknél, segélyhívásnál, jogsértés vagy bűncselekmény, tüzeset, baleset vagy más vészhelyzet bejelentésénél lehetővé teszi a nem lett nyelvű megkeresés fogadását is. Más helyzetekben, így például a börtönbüntetésüket töltők számára – akiknek kétharmadát orosznyelvűek tették ki 2000-ben – már nem engedélyezett az orosz nyelvű kommunikáció a hatóságokkal.⁵⁶

A 11. cikk (1) bekezdés értelmében a lett nyelv használatára a törvény 6. cikk (1) bekezdése értelmében kötelezett entitások rendezvényeket is csak lettül tarthatnak, ha idegen nyelvet használnak, akkor szinkrontolmacsolást kell biztosítaniuk. Kötelező az államnyelv használata a hadseregben (12. cikk), az igazságszolgáltatásban (13. cikk), míg az idegen nyelvek használatának lehetőségét a jogalkotó ebben az esetben szektorális törvényi szabályozásra hagyja. A 14. cikk biztosítja a lett nyelvű oktatáshoz való jogot, a szabályozás részleteit az oktatásról rendelkező törvényekre hagyta.

⁵⁴ A *latgal* helyzetét végül a történelmi lett területek kulturális identitásának megőrzéséről és fejlesztéséről szóló, 2021-ben elfogadott törvény rendezte, amely az állami politikai vezetés hozzáállásának egyértelmű pozitívá válását jelezte. Ennek kapcsán ld. Manzinger (2024a) i. m. 30–31.

⁵⁵ Fourth Opinion on Latvia, Advisory Committee on the Framework Convention for the Protection of National Minorities, adopted on 9 October 2023, ACFC/OP/IV(2023)I [a továbbiakban Fourth Opinion (2023)] 8. pont, <https://rm.coe.int/4th-op-latvia-en/1680ae98f6>

⁵⁶ Nils Muižnieks– Anhelita Kamenska– Ieva Leimane– Sandra Garsvāne: *Human Rights in Latvia in 2000*. Riga, Latvian Centre for Human Rights and Ethnic Studies, 2001. 41. <https://tinyurl.com/ywazjsed> Amíg Rigában visszaküldte az orosz nyelvű kérvényeket a helyi hatóság, addig Daugavpilsben az önkormányzat fordítót alkalmazott a lett nyelvi változat elkészítésére. Uo.

A média nyelvhasználatát szintén szektorális jogszabály hatáskörébe utalta a jogalkotó (16. cikk), míg a nyilvánosan bemutatott filmek esetében kötelezővé tette államnyelvi szinkronizálást vagy feliratozást, de a feliratozás esetében a főhelyre a lettnek kell kerülnie, ami nem lehetett kisebb, mint az esetleges más, idegennyelvű felirat (17. cikk).

A 18. cikk értelmében a helynevek lettül használandók Lettorszáiban, az intézmények, vállalatok lett nevet viselnek, a (4) bekezdés azonban megengedte a lív nyelvű párhuzamos feliratozást is. A 19. cikk kötelezővé tette a személynevek lett helyesírásnak megfelelő használatát – az orosz és más kisebbségi etnikumúak számára is – a (2) bekezdés annyit engedett meg, hogy a születési anyakönyvi kivonatban, illetve útlevelemben a nevek „történeti” formájának – vagyis anyanyelv szerinti írásmódjának – megfelelő, latinbetűs átíratban is feltüntethetők. A 20. cikk a bélyegzőket, iratokat kizárólag államnyelven lehet készíteni és használni – kivéve az államnyelvű és idegennyelvű változatok párhuzamos használatát lehetővé tevő kormánydöntés esetét, de ebben az esetben is a lett változatnak kell főhelyen állnia és méretben nem lehet kisebb az idegennyelvű változatnál. A 21. cikk értelmében a 6. cikk (1) bekezdés alapján lett nyelvhasználatra kötelezett intézmények köz számára szóló tájékoztatása kizárólag lettül történhet, illetve bármilyen más, a köz számára szóló információt, hirdetményt, üzenetet vagy iránymutatást államnyelven kell megjeleníteni, kivéve, ha a kormány engedélyt ad idegen nyelv párhuzamos használatára; a lett ebben az esetben is főhelyen kell álljon és méretében nem lehet kisebb az idegennyelvű változatnál. E rendelkezéseknek az a gyakorlati következménye, hogy még az olyan, a mindennapokban teljesen orosz nyelvű városokban is, mint Daugavpils, minden felirat kizárólag lettül jelenik meg.

Az államnyelvtörvény megszületését üdvözölte az EBESZ Kisebbségügyi Főbiztosa, hangsúlyozva, hogy a törvény lényegében összhangban van Lettország vállalt nemzetközi kötelezettségeivel,⁵⁷ azonban sürgette a lett nyelv magánszférában történő használatára vonatkozó előírások indokoltságához köthető mértékűre csökkentését, illetve a nemzeti kisebbségek védelméről szóló Európa tanácsi keretegyezmény ratifikációját.⁵⁸ A főbiztos által sürgetett lista 2000 novemberében készült el, a keretegyezményt pedig – hosszas halogatás után – 2005-ben ratifikálta Riga. A kihirdetésről szóló törvény⁵⁹ több szakaszában, így a keretegyezmény 10. cikkében szabályozott kisebbségi nyelvhasználat, illetve a 11. cikkben szabályozott személy- és helységnevek kisebbségi írásmód szerinti használatának és megjelenítésének jogai vonatkozásában mutatott vissza az államnyelvtörvényre, egyértelműsítve, hogy a keretegyezményben szereplő jogok a gyakorlatban csak akkor használhatók, ha az nem ellentétes az államnyelvtörvény rendelkezéseivel. Ez összhangban van a lett személynevek írásmódjáról és használatáról szóló 2004. évi 114. sz. törvény⁶⁰ előírásaival is. E jogszabály

⁵⁷ High Commissioner welcomes State Language Law in Latvia. *OSCE*, 1999. december 9. <https://www.osce.org/hcnm/52273>

⁵⁸ Statement regarding the adoption of regulations implementing the Latvian State Language Law. *OSCE*, 2000. augusztus 31. <https://www.osce.org/hcnm/52843>

⁵⁹ Likumi par Vispārējo konvenciju par nacionālo minoritāšu aizsardzību (2005. május 31.), <https://m.likumi.lv/doc.php?id=109252>

⁶⁰ A jogszabály szövege elérhető: Noteikumi par personvārdu rakstību un lietošanu latviešu valodā, kā arī to identifikāciju, <http://www.vvk.lv/index.php?sadala=134&id=687>

szerint – a hivatalos indokolás alapján azért, hogy a hatóságok mindenkit egyértelműen tudjanak azonosítani – az idegen eredetű személynevek latin betűkkel és a lett helyesírás szerint írandók. Ezek a szabályok gyakorlatilag tehát azt okozták, hogy a kisebbségi orosz nyelvet kizárták a közzsférából,⁶¹ míg a latin betűs írásmód és a lett karakterek kötelező használatára vonatkozó előírások révén nem lehet biztosítani a személynevek anyanyelvi használatának jogát.

6. A lett nyelv védelmének és fejlesztésének intézményi struktúrája

A nyelvi kérdésekkel egy komplex intézményrendszer foglalkozik, amelynek vannak hatósági és tudományos intézményi szereplői is. A Lett Tudományos Akadémián belül már a szovjet megszállás éveiben létrejött a Terminológiai Bizottság, amelynek feladata a lett nyelv fejlesztése volt, ez a testület a mai napig létezik. A munkájának segítségével hozták létre 1996-ban a Fordító és Terminológiai Központot (Tulkošanas un Terminolģijas Centrs, TTC), amely egyrészt az EU-s jogforrásokat, illetve a NATO dokumentumait fordított lettre, valamint a lett jogszabályokat EU-s nyelvekre, másrészt pedig a nyelvi harmonizációért és egységesítésért felelt – ez utóbbi funkció volt az, amiben a Terminológiai Bizottság munkáját segítette javaslataival. A TTC-t 2009-ben olvasztották be az államnyelvtörvény végrehajtását felügyelő Államnyelvi Központba (VVC), így ez utóbbi tesz azóta terminológiai javaslatokat a Terminológiai Bizottságnak, amelynek a dokumentumait végül a kormány hagyja jóvá.⁶² A Terminológiai Bizottság munkáját a nyelvészek erősen konzervatív megközelítése miatt nem minden esetben kíséri a társadalom támogatása.⁶³

Az Igazságügyi Minisztérium alárendeltségében működő VVC kiemelt feladata az államnyelvtörvény 26. cikke értelmében a törvény végrehajtásának ellenőrzése, aminek keretében bírságot is rendelkezik, amivel nem ritkán él is. A VVC emellett – a 2020. szeptember 2-i, 550. sz. kormányhatározattal elfogadott szabályzata alapján – az államnyelv beszélőinek jogait és érdekeit védi, előmozdítja a lett nyelv használatát a legkülönbözőbb szférákban, elősegíti a lett Európai Unió intézményekben történő minél teljesebb körű használatát.⁶⁴ Feladatának ellátása érdekében egyebek között jogosult törvényben meghatározott esetekben térítésmentesen felvilágosítást kérni magánszemélyektől, előzetes engedély, belépődíj vagy más korlátozás nélkül ellenőrzési tevékenységet végezni állami és önkormányzati intézményekben, kereskedelmi vállalatoknál és szervezeteknél, felszólítani a törvényellenes magatartás vagy gyakorlat megszüntetésére, beidézni személyeket, illetve tanácsadó véleményt adni. A VVC munkatársainak és igazgatóhelyettesének döntéseivel szemben az igazgató-

⁶¹ Michele E.Commercio: Exit in the Near Abroad – The Russian Minorities in Latvia and Kyrgyzstan. *Problems of Post-Communism*, 2004/6. 25.

⁶² 1999-es államnyelvtörvény 22. cikk (2) bekezdés.

⁶³ Kruk (2011) i. m. 452–458. Az egyik legnagyobb felháborodást kiváltó nyelvújítás-kezdemenyezésük 2004-ben az euró lett megnevezésének megváltoztatására irányult, amit a társadalmi tiltakozások okán kénytelenek voltak visszavonni. Uo.

⁶⁴ A szabályzat elérhető: Valsts valodas centra nolikums. <https://tinyurl.com/hfk95hwf>

hoz, az igazgató döntésével szemben a kerületi (városi) bírósághoz lehet fellebbezést benyújtani, míg az igazgató adminisztratív intézkedéseivel szemben az Igazságügyi Minisztériumtól lehet jogorvoslatot kérni, a minisztérium döntésével szemben pedig bíróságtól.

Az akkori államelnök, Vaira Viķe-Freiberga kezdeményezésére hozták létre 2002. január 16-án az Állami Nyelvi Bizottságot (Valsts Valodas Komisija), amelynek az lett a feladata, hogy „vizsgálja a nemzeti nyelv helyzetét Lettországhban, és javaslatokat dolgozzon ki a lett mint nemzeti nyelv pozícióinak erősítésére, illetve a nyelvnek a fejlesztésére”.⁶⁵ A bizottság tagjai nyelvészek, kulturális, tudományos és oktatási szakemberek, illetve a társadalom más részeinek képviselői. A Bizottság mellett működik a Lett Állami Nyelvoktatási Ügynökség (Latviešu valodas apguves valsts aģentūra, LVAVA), amelyet az 1995 óta működő Lett állami nyelvtanulási program (Latviešu valodas apguves valsts programma) utódaként 2004-ben hoztak létre azzal a céllal, hogy a lett nyelvnek a nemzeti „kisebbségek” – elsősorban az oroszok – körében történő oktatását megszervezze, illetve hatékonyabbá tegye. A nyelvoktatáson kívül tanárképzéssel, tananyagok és módszertani segédanyagok kidolgozásával, valamint társadalmi kapcsolatok kiépítésével foglalkozik.

Az állami intézmények munkáját jelentős mértékben meghatározzák a kormányzat által készített programok. Az első társadalmi integrációs program – nemzetközi nyomásra, az EU-csatlakozás előkészítéseként – 2001-ben került elfogadásra. A dokumentum megfogalmazta azt a máig is érvényes alapvetést, hogy a társadalmi integráció útja a lett kulturális értékek elfogadása, a lett nyelv használata és az állam iránti lojalitás, a nem-lettek lett nyelvhasználata pedig a lettek bizalmának elnyerését segítő feltétel.⁶⁶ A következő dokumentumok, így a most irányadó, *A hivatalos nyelvpolitika alapvetései 2021–2027 c. kormányzati stratégia* is ezt a célt tűzték ki.

7. Nyelvi jogok a politikai térben

Habár természetesen a minisztériumok mellett működő, fentebb bemutatott intézmények is politikai térben léteznek, két olyan kérdésre érdemes kitérni, amelyek Lettország számára kihívást jelentenek: az egyik egy 2012-es népszavazás, a másik pedig az Emberi Jogok Európai Bíróságának (EJEB) 2023. május 2-i ítélete a Mestan kontra Bulgária ügyben.

Lettországban 2012. február 18-án társadalmi kezdeményezésre országos népszavazást tartottak az orosz nyelv második hivatalos nyelvvé nyilvánításáról. Ez a gyakorlatban azt jelentette volna, hogy – hasonlóan a letthez – az orosz nyelv védelmére is fel kellett volna esküdniük a parlamenti képviselőknek, az szabadon használhatóvá vált volna a parlament és az önkormányzatok munkájában, illetve az állami és önkormányzati intézményekkel folytatott kommunikációban. Lett értékelés szerint a szimbolikus jelentőségén túl a kezdeményezés azt célozta volna, hogy a lettországi orosznyelvűek

⁶⁵ Valsts valodas komisija, <https://tinyurl.com/mpd7ms97>

⁶⁶ Kruk (2022) i. m. 73.

megőrizhessék egynyelvűségüket és ne legyenek a lett megtanulására kényszerítve.⁶⁷ A szavazáson a jogosultak 71,2%-a vett részt, ennél magasabb részvétel csak az 1991-es függetlenségi népszavazáson volt úgy, hogy akkor valamennyi lakos szavazhatott, míg 2012-ben a lakosság mintegy egyhatedét kitevő – döntően orosz nyelvű – nem-állampolgárok nem.

Az orosz második hivatalos nyelvvé nyilvánítását a megjelentek 24,9%-a támogatta; Rigában 36% volt az igenek aránya, a délkeleti Latgale régióban 56%, míg az ország második legnagyobb városában, Daugavpils-ben meghaladta a 85%-ot. Összességében azonban a jogosultak több mint 53%-a és a szavazáson résztvevők 74,1%-a elutasította a kezdeményezést.⁶⁸ A népszavazást követően módosították az alkotmány preambulumát, amibe belekerült, hogy az állam egyetlen nyelve a lett, a lett identitás pedig a lett és lív hagyományokra, a lett népi bölcsességre, a lett nyelvre, valamint az egyetemes emberi és keresztény értékekre épül. Ezzel összehasonlítva némi aránytalanságot mutat az alkotmány változatlan 114. cikke, amely deklarálja a nemzetiségek nyelvük, kultúrájuk és etnikai örökségük megőrzéséhez való jogát.⁶⁹

A másik politikai szempontból kiemelt kérdés a választási kampányokról szóló törvény 2023. júniusi módosítása, ami megtiltotta a letten kívül más nyelv használatát a kampányolás során – a helyi önkormányzati és EP-választások kivételével, ahol az EU-hivatalos nyelvek használhatók maradtak. A keretegyezmény tanácsadó bizottsága 2023-ban ezt egyszerre nevezte az anyanyelvhasználat 9. cikkben és a közéletben való hatékony részvétel 15. cikkben biztosított jogainak megsértésének,⁷⁰ és emlékeztetett, hogy az EJEB 2023. május 2-i Mestan kontra Bulgária ügyben⁷¹ elmarasztalta a balkáni államot, mert a 2013-as választási kampányban a török nyelven is kampányoló Lütfi Mestan pártvezért Sliven régió kormányzója megbírságolta a hivatalos nyelv – alkotmányban rögzített – kötelező használatának megsértéséért. Az EJEB a hivatalos nyelvtől eltérő nyelv használatának tilalmát az Emberi Jogok Európai Egyezménye 10. cikkében biztosított véleménynyilvánítás szabadságához fűződő jog megsértésének minősítette, a döntés pedig a hasonlóság miatt intő példa Lettország számára.

8. A nyelvi jogok gyakorlata Lettországban

Mára a lettországi társadalom kétnyelvűvé vált: mivel az orosz nyelv – még – második idegen nyelvként szinte minden lett iskolában tanított, így a mai fiatal generációk orosz nyelvtudása is jó, miközben a lettországi orosz jelentős része is beszél ma már lettül. Ebben jelentős segítséget jelent az, hogy a lettországi oroszok 40%-a él szórványban,

⁶⁷ Druviete (2002) i. m. 3.

⁶⁸ Az adatok régiókra, kerületekre, településekre és szavazókörökre bontva elérhetőek: 2012. gada 18. februāra tautas nobalsošana par likumprojekta „Grozījumi Latvijas Republikas Satversmē” pieņemšanu. Rezultāti, <http://www.tn2012.cvk.lv/report-results.html>

⁶⁹ Alijeva (2017) i. m. 512–513.

⁷⁰ Fourth Opinion (2023) i. m. 184. pont.

⁷¹ Affaire Mestan c. Bulgarie (Requête no 24108/15), 2023. május 2., <https://tinyurl.com/3wc38ucv>

míg 48%-a lett–oroszlinteretnikus közegben és csak 12%-uk, a Daugavpilsben 85%-ot kitevő orosz nyelvűek élnek ma még feltörhetetlen nyelvi zárványban.⁷²

Ahogy arra a fentebb hivatkozott Tanácsadó Bizottság folyamatosan rámutat, Lettország vonatkozó szabályozásai nincsenek összhangban kevés vállalt nemzetközi kötelezettségével, így elsősorban a keretegyezményben foglaltakkal. Állami vagy önkormányzati intézmények előtt – az államnyelvtörvény 10. cikk (2) bekezdésben felsorolt kivételektől eltekintve –, illetve azok belső működésében csak a lett nyelv használható. Ez a mindennapokban azt jelenti, hogy – ezt a szerző személyesen tapasztalta – intézményekben nem lett nyelvű kommunikációt csak idegenek elől elvonulva, zárt ajtók mögött folytatnak, illetve angol nyelvű beszélgetés során olyan hivatalos személyek is lettül beszélnek a tolmáchoz, ahol mind az érintett, mind a tolmács jobban beszél oroszul. A törvény megfélemlítő hatása másfajta önkorlátozást is eredményez: korábban kifejezett törvényi tiltás hiányában sok közintézmény és helyi önkormányzat tartott fenn orosz nyelvű honlapot is a lett nyelvű mellett, ezeket azonban önkéntes alapon elkezdtek felszámolni vagy egyszerűen csak nem frissítették a 2022. februári orosz agressziót követően.⁷³

A nem lett – ideértve a latgal – vagy lív nyelvű földrajzi nevek kihelyezése továbbra is tilos, a személyneveket pedig a hivatalos dokumentumokban latin betűkkel, a lett helyesírás szabályainak megfelelő átírással kell feltüntetni, ami a gyakorlatban azt jelenti, hogy a férfinevek esetében egy „s”, a női nevekében pedig egy „a” vagy „e” betű végződést adnak hozzá mind a család-, mind a keresztnévhez. Ez – a Tanácsadó Bizottság legfrissebb véleményét idézve – nem felel meg a keretegyezmény 11. cikkében biztosított jogoknak és sérti az egyének személyes identitásához, továbbá személyi méltóságához fűződő jogait.⁷⁴ Szintén nehezen érthető helyzetet eredményezett az, hogy a COVID-járvány idején nagyon szűk volt a nem lett nyelvű tájékoztatás lehetősége, az államnyelvtörvény fent idézett előírása alapján például az oltásokkal kapcsolatos információkat kizárólag lettül lehetett megosztani.⁷⁵ Hasonló helyzet merült fel az itt nem vizsgált oktatás vonatkozásában is a járvány idején. 2021-ben ugyanis az ombudsman az Oktatási és Tudományos Minisztériumhoz fordult azért, hogy biztosítsa az egyenlő hozzáférést az online elérhető távoktatási anyagokhoz a kisebbségi tannyelven tanuló diákok számára is, de a helyzet megnyugtató kezelése nem történt meg.⁷⁶

A média területén – az általános államnyelvtörvényi előírások mellett – szektorális jogszabályokban adják a jogi keretet. A média esetében a legnagyobb kihívás Lettország számára az oroszországi médiafogyasztás elterjedtsége: 2022-ben a lakosság

⁷² Martin Ehala: After Status Reversal: The Use of Titular Languages and Russian in the Baltic Countries. *Harvard Ukrainian Studies*, 2017–2018/1–4. 476–477.

⁷³ Fourth Opinion (2023) i. m. 102. pont.

⁷⁴ Uo. 104., 110–112. pontok.

⁷⁵ Uo. 103. pont. A COVID-járvány során a kisebbségi nyelvhasználattal más országban is voltak problémák, több esetben a hivatalos állam- vagy regionális nyelvre vonatkozó előírások miatt. Ezzel kapcsolatosan ld. Manzinger Krisztián: Nyelvi jogok a koronavírus idején – Egyes európai államok gyakorlata a COVID-19 járvány során 2020 első félévében. *Glossa Iuridica – Jog és vírus 2020*, VII. különszám, 235–247.

⁷⁶ Fourth Opinion (2023) i. m. 35. pont.

63%-a, azon belül a lett nemzetiségűek 51%-a, míg a kisebbségek 82%-a fogyasztott oroszországi sajtóterméket, sőt, a lettországi orosz nyelvűek 46%-a egyáltalán nem fogyasztott lett nyelvű médiát.⁷⁷ Az oroszországi médiafogyasztással és az orosz nyelvű műsorszórással szemben az állam hosszabb ideje igyekszik fellépni: a 2010-es elektronikus médiáról szóló törvény⁷⁸ kötelezi az országos és regionális közszolgálati és kereskedelmi elektronikus műsorszórókat, hogy legalább 65%-ban államnyelven sugározzanak, míg a közszolgálati műsorszórásban, habár az orosz nyelv továbbra is jelen van – az ukrán és a lengyel mellett –, a műsoridő jelentősen csökkent az elmúlt években: 1.169 órától 318 órára 2018-2022 között.⁷⁹ Az oroszországi Rossiya RTR sugárzását a Lett Médiatanács 2014-ben, 2016-ban és 2021-ben is felfüggesztette Lettország területén gyűlöletbeszédre és etnikai feszültségek szítására hivatkozva, majd a elektronikus média-törvény nemzetbiztonsági okokra visszavezethető 2022-es módosításával meg is szüntette azt – összesen 130 oroszországi TV-csatornával és 169 weboldallal együtt,⁸⁰ mindazonáltal az ilyen korlátozások például VPN-nel kikerülhetők.

A 2023-ban elfogadott, az *Információs Tér Biztonságáról és a Stratégiai Kommunikációról* szóló, 2023-2027-es időintervallumú nemzeti stratégia⁸¹ kiindulópontja az, hogy mivel Oroszország a legfontosabb nemzetközi kihívás Lettország számára az információs térben, az orosz média célja pedig a lettországi kisebbségek befolyásolása és ezzel a lett társadalom megosztása, ez ellen az országnak saját védelme érdekében fel kell lépnie.⁸² A kormányzati célkitűzés ebben a tekintetben a dezinformációs oroszországi média – 2022-ben már elvileg megvalósult – elérhetetlenné tételét követően rövid távon a meglévő lettországi média felhasználásával elérni e célcsoportokat – de úgy, hogy az nem eredményezi az állami forrásokból finanszírozott kisebbségi nyelvű média növekedését –, közép- és hosszútávon pedig a lett nyelvű tartalom fogyasztásának emelésével bevonni az érintetteket az európai média és információs térbe, ami egyúttal a lett társadalom államnyelvi alapú konszolidációját is előmozdítja.⁸³ A stratégia alapján már 2023 tavaszán tárgyalt a lett törvényhozás arról, hogy 2025 végéig megszüntetik a nem lett, nem EU-s és nem tagjelölt országok nyelvén való műsorszórás lehetőségét – vagyis elsősorban az orosz. A tervet a keretegyezmény Tanácsadó Bizottsága élesen kritizálta, hangsúlyozva, hogy nemcsak az anyanyelvű médiához hozzáférés jogát védi a keretegyezmény, hanem azt is tartalmazza, hogy a nyílt és pluralista médianak lehetőségét kell biztosítani a kisebbségeket foglalkoztató kérdések megvitatására, ezért nem az orosz nyelvű médiatartalom csökkentése lenne kívánatos, hanem inkább annak

⁷⁷ Uo. 93. pont.

⁷⁸ Elektronisko plašsaziņas līdzekļu likums. <https://tinyurl.com/9yzd7zpv>

⁷⁹ Fourth Opinion (2023) i. m. 90. pont.

⁸⁰ Uo. 94. pont.

⁸¹ The National Concept on Strategic Communication and Security of the Information Space 2023–2027. <https://www.mk.gov.lv/en/media/15446/download?attachment>

⁸² Uo. 8.

⁸³ Uo. 18.

növelése, és ez – az állam által ellenőrzött tartalmat közvetítő média – lehetne az útja annak, hogy a társadalmi integrációt az országban erősítsék.⁸⁴

A nyelvi jogok kérdése a szankcionálás lehetősége miatt tisztségviselőkkel szembeni fellépésre is alkalmas. Rihards Eigims korábbi daugavpils polgármestert például 2017 októberében bírságolták meg az államnyelv nem megfelelő tudásáért, illetve kötelezték arra, hogy hat hónapon belül ismételt nyelvvizsgát tegyen.⁸⁵ Az ET keretegyezményének alkalmazását vizsgáló Tanácsadó Bizottság ezt 2018-ban a demokratikus folyamatokba való nem megfelelő beavatkozásnak minősítette és felhívta a lett hatóságokat, hogy keressenek más módot az államnyelv ismeretének előmozdítására.⁸⁶ Máshogy alakult viszont Ivans Baranovs Balvi városi tanácsos ügye, akit a lett Legfelső Bíróság 2016. december 13-i döntésével – elégtelen államnyelvtudás okán – megfosztott mandátumától.⁸⁷ Baranovs az ENSZ Emberi Jogi Bizottságához fordult⁸⁸ a Polgári és Politikai Jogok Nemzetközi Egyezségokmánya 25. cikkében garantált jogai⁸⁹ megsértésére hivatkozással. A testület előtt Lettország azzal érvelt, hogy Baranovs többször is megtagadta a megjelenést az Államnyelvi Központ előtt, így nem lehetett igazolni, hogy birtokában van a feladata ellátásához szükséges államnyelvtudásnak. Az Emberi Jogi Bizottság 2022. február 10-i döntése 8.4. pontjában az államnyelv megfelelő szintű ismeretét olyan megfelelő indoknak nevezte, amely lehetőséget ad a 25. cikkben foglalt jogok korlátozására, amennyiben kellően világosan megfogalmazott szabályon alapul és nem valósít meg diszkriminációt, így Baranovs beadványát elutasította.⁹⁰

⁸⁴ Fourth Opinion (2023) i. m. 97. pont.

⁸⁵ Eigims-et 2018-ban korrupciós vádak miatt elmozdították a polgármesteri székből, majd 2022 áprilisában hároméves börtönbüntetésre ítélték korrupció miatt.

⁸⁶ Advisory Committee on the Framework Convention for the Protection of National Minorities, Third Opinion on Latvia adopted on 23 February 2018, ACFC/OP/III(2018)001REV [a továbbiakban Third Opinion (2018)] 167. pont, 47. o., <https://tinyurl.com/ymtdu6k5>

⁸⁷ 2017 júniusában a helyiek újraválasztották. Third Opinion (2018) i. m. 167. pont, 47. o.

⁸⁸ Views adopted by the Committee under article 5 (4) of the Optional Protocol, concerning communication No. 3021/2017, CCPR/C/133/D/3021/2017, *United Nations Human Rights Committee*, <https://tinyurl.com/cfzdrx2s>

⁸⁹ 25. Cikk

Minden állampolgárnak a 2. Cikkben említett megkülönböztetések, illetőleg ésszerűtlen korlátozások nélkül joga és lehetősége van arra, hogy

- a) a közügyek vitelében közvetlenül vagy szabadon választott képviselők útján részt vegyen;
- b) szavazzon és megválaszthatassák az általános és egyenlő választójog alapján, titkos szavazással tartott olyan valódi és rendszeres választásokon, amelyek biztosítják a választók akaratának szabad kifejezését;
- c) az egyenlőség általános feltételei alapján hazájában közhivatali tisztséget viselhessen.

1976. évi 8. törvényerejű rendelet az Egyesült Nemzetek Közgyűlése XXI. ülészakán, 1966. december 16-án elfogadott Polgári és Politikai Jogok Nemzetközi Egyezségokmánya kihirdetéséről.

⁹⁰ A döntéshez különvéleményt csatolt Furuya Shuichi bizottsági tag, aki úgy vélte, hogy habár az államnyelv ismerete fontos szempont, de csak egy a több szempont közül, amelyek alapján valaki alkalmas egy pozíció betöltésére. Mivel az adott esetben (többször is újra)választott tisztviselőről van szó, így a választók meglátása szerint a feladatának ellátására alkalmas, így a mandátumától való bírósági megfosztás nem objektív és nem indokolható.

A lett nyelvpolitika nemzetközi bírósági eljárásokban is megjelent, így például a Podkolzina kontra Lettország ügyben⁹¹ az EJEB megállapította az Emberi Jogok Európai Egyezményéhez csatolt Első kiegészítő jegyzőkönyv szabad választásokhoz való jogról szóló 3. cikkének a megsértését, de az egyezmény 14. cikkében foglalt megkülönböztetés tilalmának megsértését nem. Az adott esetben az történt, hogy az 1998-as lettországi választás egyik jelöltjének munkahelyén két egymást követő nap is megjelent bejelentés nélkül az Államnyelvi Központ vizsgáztatója, aki először szóbeli, majd írásbeli lett nyelvtudása bizonyítására kötelezte Ingrida Podkolzina képviselőjelöltet, amelyen az megbukott, így törlésre került a jelöltek listájáról. Az ítéletében az EJEB emlékeztetett a lett állam álláspontjára, amely az 1993-as Saecima tapasztalatai alapján – amelynek egyes tagjai a munkát akadályozó mértékben nem beszélték a lett nyelvet – a parlamenti képviselőktől elvárta a folyékony lett nyelvtudást, így nem állapított meg diszkriminációt e tekintetben, azonban a hiányos nyelvtudás miatti törlést a képviselőjelölti listáról aránytalan intézkedésnek minősítette és az Első kiegészítő jegyzőkönyv 3. cikkének megsértése miatt elmarasztalta Lettországot.

A nyelvtudás továbbra is feltétele a lett állampolgárság megszerzésének, a honosítás iránt azonban gyenge az érdeklődés. 2022-ben azonban összekapcsolták a nyelvtudás kérdését az állandó tartózkodási engedély birtoklásának jogával az orosz állampolgárok vonatkozásában: az 1990-es években a lett állampolgárságból kirekesztettek egy kis része orosz állampolgárságot vett fel és Lettország területén maradt. Őket Riga kiemelt biztonsági kockázatként kezel, különösen az Ukrajna elleni orosz támadás kontextusában. Az állandó tartózkodási engedéllyel rendelkező orosz állampolgároktól 2022 végén visszavonták ezen engedélyeket és az új engedélyek kibocsátását A2-es szintű lett nyelvtudás igazolásához kötötték.⁹²

Az egyébként is rendkívül szűkre szabott nyelvi jogok továbbra is politikai mezőben vannak: a 2022. őszi választásokat megelőzően az államnyelvtörvény további – magánszférát érintő – módosítása is felmerült, így a nem lett vagy EU-nyelvek használatának megtiltása a munkahelyeken, ez azonban a választásokat követően lekerült a napirendről. A keretegyezmény Tanácsadó Bizottsága mindazonáltal hangsúlyozta 2023-as jelentésében, hogy az állam akkor jogosult valamely nyelvek használatát korlátozni, ha az közbiztonsághoz, közegészségügyhöz kapcsolódik, fogyasztói vagy munkajogi érdekek, illetve munkabiztonsági okok indokolják, a korlátozásnak azonban arányosnak és a szükségesség elvének megfelelőnek kell lennie; ennek pedig sem a hatályos, sem a korábban tervezett további szigorítások nem felelnek meg.⁹³ Megtörtént viszont a nyelvi jogok szűkítése az oktatás területén: egy 2022-es törvénymódosítással 2023. szeptember 1-jével megkezdődött – felmenő rendszerben 2025. szeptember 1-ig bezárólag – az orosz nyelvű oktatás és intézményrendszerének felszámolása, az ENSZ vezető tisztviselőinek tiltakozása mellett.⁹⁴ Ezzel a fejleménnyel kapcsolatosan érdemes utalni az EJEB Valiullina és mások kontra Lettország ügyben 2023. szeptember 14-én hozott

⁹¹ Podkolzina kontra Lettország no. 46726/99 (ECHR, 2002. április 9.)

⁹² Manzinger (2024b) i. m. 55.

⁹³ Fourth Opinion (2023) i. m. 105–109. pontok.

⁹⁴ Manzinger (2024c) i. m. 83–84.

ítéletére is, amelyben a bíróság a lett használatának ötven évet meghaladó elnyomására hivatkozva legitimnek találta az államnyelv használatának erősítésére irányuló állami törekvéseket. Az EJEB ezen felül azt is hangsúlyozta, hogy mivel Lettország egységes oktatási rendszert működtet, az országban pedig az egyetlen hivatalos nyelv a lett, ezért az orosz nyelv használatának csökkentése nem jelenti Lettország vállalt nemzetközi kötelezettségeinek megsértését.⁹⁵

9. Összegzés

A középső balti állam 1918–1920-ban elnyert függetlenségével első ízben szerzett lehetőséget arra, hogy a lett nyelvet hivatalos használatba emelje. Ezzel olyan több évszázados időszakot zártak le – az 1940–1991 közötti szovjet megszállás okán először csak ideiglenesen –, amelyben a térségben más nyelvek voltak dominánsak: a német, illetve a lengyel, majd az orosz. Az első függetlenség első másfél évtizedében egy hivatalos többnyelvűsége épülő állami politika valósult meg a lett dominanciájával, majd ezt az 1934-es puccs után a lett államnyelv kizárólagossága követte. Az 1940-es, majd 1944-es szovjet megszállás visszahozta az orosz dominanciáját a lettországi valóságba, a következő évtizedek pedig a lett nyelvű lakosság lett–orosz kétnyelvűvé válását, az orosz és más kisebbségi nyelvű lakosság esetében pedig döntően az orosz nyelv megtartását vagy dominánssá válását hozták, amely a lett nyelvet defenzívába kényszerítette.

A függetlenség visszaállítását követően 1992-ben jelentősen módosított formában léptették hatályba az 1989. évi nyelvtörvényt, amely szakított a lett–orosz kétnyelvűséggel és immár csak egyetlen hivatalos nyelvet, a lettet ismerte el. A máig tartó folyamatos jogalkotás a lett nyelv pozícióinak erősítéséről szól, amelynek már megvannak az alkotmányos alapjai, illetve a szektorális jogszabályai. Az alkotmányos alapelvek mellett az 1999. évi államnyelvtörvény tartalmazza a legfontosabb rendelkezéseket, amelyeket kiegészítenek a médiára, oktatásra, névhasználatra és más területekre vonatkozó szabályok. Összességében a szabályozás célja az orosz nyelv pozícióinak csökkentése lehetőség szerint szinte minden szabályozható élethelyzetre a kötelező lett nyelvhasználat előírásával és a jogsértés szankcionálásával. Az elmúlt évtizedekben kiépült a szabályozásnak keretet biztosító intézményrendszer is, amely részben a lett nyelv fejlesztését, részben pedig az államnyelvtörvény végrehajtását ellenőrzi, illetve a törvénysértést szankcionálja. A megteremtett jogi keret teljesen egyoldalú, ezt a nemzetközi kritikák – elsősorban a keretegyezmény Tanácsadó Bizottsága – visszatérően hangsúlyozzák, mert a lett nyelvtudást igen, de a társadalmi kohéziót biztosan nem erősíti. Igaz ez annak ellenére is, hogy a lettországi orosznyelvűek jelentős része valóban kevés tiszteletet mutat a lett nyelv és kultúra iránt, ami a honosítás iránti alacsony érdeklődés mellett a mindennapi életben is tapasztalható.

Ami magyar szempontból kifejezetten fontos, az az ukrainai szabályozás hasonlósága. Ukrajna elmúlt évtizedben zajló erős nemzetállamépítő törekvései jelentős mértékben építenek a lettországi szabályozásra és tapasztalatokra – ennek alátámasztására

⁹⁵ Valiullina and Others v. Latvia (applications nos. 56928/19, 7306/20 és 11937/20). ECHR 256 (2023) 14.09.2023 ; Manzinger (2024c) i. m. 81.

a hasonló, posztszovjet létet és az orosz nyelv dominanciájának erősítésére irányuló korábbi orosz, majd szovjet állami törekvéseket hozzák fel. A lettországi és az ukrainai helyzet kapcsán azonban van egy lényegi különbség, ami elsősorban, de nem kizárólagosan a kisebb ukrainai nemzeti közösségeket érinti: a magyar, román és bolgár lakosság, de a krími orosz is határváltozások révén kerültek Ukrajna fennhatósága alá, vagyis nyelvük nem korábbi állami nyelvpolitikai törekvések miatt van jelen az ország területén, hanem azért, mert történelmi események révén Ukrajna olyan területre terjesztette ki fennhatóságát, ahol ezek a nemzetiségek hagyományosan élnek. Amíg az államnyelv erősítésére irányuló törekvések Lettországból kifejezetten arra irányulnak, hogy a szovjet időszakban a lettekre erőltetett kétnyelvűséget eltolják az országuk területére zömmel a szovjet időszakban betelepített lettországi kisebbségek kétnyelvűségének irányába – még ha a módszereket akár jogos kritikák is érik –, addig Ukrajna a lett mintát arra használja, hogy az ország területén hagyományosan élő nemzetiségek nyelvhasználatát korlátozza – ebben pedig sokkal inkább hasonlít a szlovák államnyelvi szabályozáshoz, mint a lettchez.

DISSERTATIONES

BEVEZETŐ A MESTERSÉGES INTELLIGENCIA
ÉS A MÉDIA KAPCSOLATÁT TÁRGYALÓ
TANULMÁNYOKHOZ

A mesterséges intelligencia (MI) és a média kapcsolata napjaink egyik leggyorsabban fejlődő és egyben legösszetettebb kutatási területe, amely sajátos kihívásokkal szembesíti a tudományos elemzést végző kutatókat. A technológiai fejlődés üteme olyan gyors, hogy a hosszú távú és stabil állítások megfogalmazása szinte lehetetlen vállalkozásnak bizonyul. A mesterséges intelligencia alkalmazásának technológiai, szabályozási és gyakorlati aspektusai napról napra változnak, ami folyamatos revízióra készíti a tudományos közösséget. E rendkívüli dinamikusság ellenére is nélkülözhetetlen a mesterséges intelligencia média területén gyakorolt hatásainak mélyreható elemzése, mivel ezek a hatások alapjaiban formálják át nemcsak a média működését és szerkezetét, hanem szélesebb értelemben vett társadalmi nyilvánosságot is. A mesterséges intelligencia alapú technológiák alkalmazása révén alapvetően változik meg a média tartalomelőállításának módja, a tartalom terjesztése, valamint annak fogyasztása is. Ezek a változások pedig rengeteg új kérdést vetnek fel a média jelenségeit kutató tudományterületek számára.

A jelen tematikus blokkban közölt tanulmányok közös pontja az MI és a média kérdésköre, a médiatudatosságtól az algoritmusokkal kapcsolatos kérdéseken át a dezinformációig. A tudatos médiafogyasztás fogalma immár messze túlmutat a hagyományos médiaértés koncepcióján, és magában foglalja az algoritmusok működési elveinek megértését, a felhasználók kritikai gondolkodásának fejlesztését, valamint az álhírek felismerésének és kezelésének képességét. Az MI-alapú rendszerek fokozódó alkalmazása révén a médiafogyasztás egyre személyre szabottabbá válik, ami nem pusztán követi, hanem aktívan formálja is a felhasználói preferenciákat, fogyasztói viselkedést és percepciókat. Ezek a személyre szabási mechanizmusok jelentős hatással bírhatnak a közvélemény alakulására, a társadalmi diskurzus minőségére, illetve a politikai és társadalmi döntéshozatalra is. A tanulmányok másik kiemelt témája az algoritmusok elszámoltathatóságának kérdése, amely az algoritmusok vezérelte médiakörnyezetben kiemelten releváns probléma. Az algoritmusok működésének átláthatósága, ellenőrizhetősége és a döntési folyamatok megismerése technikai, etikai, jogi és társadalmi kérdéseket egyaránt felvet. A médiában alkalmazott algoritmusok nem válhatnak zárt, átláthatatlan rendszerekké, egy demokratikus társadalomban



különösen fontos, hogy a média működése átlátható és ellenőrizhető maradjon, így biztosítva a tájékozott közvélemény kialakulását, valamint a nyilvános vita szabad és nyitott jellegét.

Kiemelten fontos továbbá az algoritmusokban rejlő esetleges elfogultságok és torzítások feltárása és kritikus elemzése. Az algoritmusok működésében megjelenő torzítások ugyanis nem csupán az információ elérhetőségét, hanem a társadalmi percepciókat és viselkedésmintákat is jelentősen befolyásolhatják. Ezek a torzítások láthatatlan módon erősíthetik fel a társadalmi egyenlőtlenségeket, előítéleteket és polarizációs folyamatokat, továbbá mélyíthetik az információs és digitális szakadékokat. Ennek következtében a média által közvetített tartalom torzítása társadalmi és politikai attitűdök formálódását is befolyásolhatja, így széleskörű és messzemenő társadalmi következményekkel járhat. Az algoritmusok működésének részletes vizsgálata ezért multidiszciplináris kutatói együttműködést igényel, amely magában foglalja a társadalomtudományok, a jogtudomány, a kommunikációtudomány, valamint a technológiai tudományok közös munkáját.

A jelen tanulmányok természetesen nem tekinthetők véglegesnek vagy lezártak, tekintettel arra, hogy a mesterséges intelligencia és a média viszonyának problémaköre folyamatosan változik és új kutatási kérdéseket vet fel. A tudományos közösség számára folyamatos feladat az új ismeretek integrálása, a korábbi állítások revíziója és az új technológiai fejlesztések hatásainak elemzése. A tematikus szekció tanulmányai célul tűzik ki, hogy átfogó, alapos és sokoldalú képet adjanak a tárgyalt kérdésekről, ezáltal ösztönözve a további kutatásokat és hozzájárulva az e gyorsan változó területen folyó tudományos párbeszédhez.

Papp János Tamás
adjunktus, PPKÉ JÁK;
a lapszám vendégszerkesztője

MESTERSÉGESINTELLIGENCIA-JÁRTASSÁG

Andok Mónika*

1. Bevezetés

A tömegkommunikációs eszközöknek, a médiának a hatása régóta kutatott terület; tulajdonképpen az elektronikus tömegkommunikációs eszközök megjelenése óta. A tudományos kutatások egyre határozottabb diszciplináris fókusszal foglalkoztak a rádiózás, a mozi, majd a televíziós tartalomfogyasztás egyéni és társadalmi következményeivel, először a pszichológia és a szociológia, majd a formálódó médiatudomány keretében. A média hatásainak egyre pontosabb tudományos feltérképezését hamarosan követte egy pedagógiai célkitűzés is, miszerint a média nem kívánt, káros hatásainak mérsékléséhez szükséges a média használatának oktatása, a médiatartalmakhoz kapcsolódó kritikai attitűd, gondolkodásmód kialakítása. Ezen készségek átfogó elnevezése a médiajártasság lett, melynek angol megfelelője, a *media literacy* elnevezés árulkodik arról, hogy ezt a készségcsoportot az írott szó elolvasásának képességéből vezették le.¹ Majd a további médiatechnológiák megjelenésével a kifejezés további jelentésbővülésen ment keresztül, s ma már beleértik az információs, vizuális műveltséget is; a médiára vonatkozó elméleti tudást és a használatához köthető jártasságot. S bár a *literacy* kifejezéssel kapcsolatban kritikai érveket is felsorakoztattak a médiakutatók, a kifejezés végül is meghonosodott úgy az akadémiai szférában, mint a médiaszabályozás vonatkozásában, nemzetközi és magyar környezetben is.² Potter 2015-ben így definiálta: „A médiaműveltség azon perspektívák készlete, melyeket aktívan használunk, amikor a tömegkommunikációs rendszerekkel kapcsolatba kerülünk, és a bennünket elérő üzeneteket értelmezzük.”³ Ezen perspektívákban a kilenc-

* Habilitált egyetemi docens, Pázmány Péter Katolikus Egyetem, Bölcsész- és Társadalomtudományi Kar, Kommunikáció Tanszék. ORCID: <https://orcid.org/0000-0002-3137-1205>

¹ James W. Potter: *Médiaműveltség*. Budapest, Wolters Kluwer, 2015. 15.

² Sonia Livingstone: Media literacy and the challenge of new information and communication technologies. *The Communication Review*, Vol. 7., No. 1. (2004) 4.; Nagy Krisztina: *Műveltség, média, szabályozás: a médiaműveltség médiapolitikai jelentősége és szabályozási keretei*. Budapest, Gondolat, 2018.

³ Potter i. m. 25.



venes évek végétől kezdve megjelenik a digitális állampolgárság kérdése is. Vagyis a médiaműveltség szintje, mértéke nemcsak az egyén, de a közösség, a demokratikus működő közösség szempontjából is fontos szerepet kap. Összességében tehát azt mondhatjuk, hogy a médiaműveltségnek van egy média- és információs technológiához köthető folyamatos bővülése, reaktualizációja; egyéni, közösségi és társadalmi szintű hasznossága; deklaratív tudásból és operacionalizálható készségekből álló területe. A technológiai fejlődés azt is jelenti a felhasználók és a kutatók számára, hogy minden új kommunikációs és médiatechnológia megjelenésekor a sikeres, eredményes használatához új készségek elsajátítására van szükség. Vagyis a mesterségesintelligencia-jártasság igénye nem üres területre érkezik, hanem korábbi médiajártasságokhoz fog kapcsolódni, mind a koncepció mind a taxonómia tekintetében.

2. Médiaműveltség – mesterségesintelligencia-műveltség

A 21. században a médiaműveltség készségei segítik az embereket a média intelligens használatában, a média tartalmainak megkülönböztetésében és kritikus értékelésében. A médiaműveltség a legtöbb fejlett országban több mint negyven éve az oktatás része. A médiaműveltség kiterjed nemcsak a média tudatos és kritikus használatára, a hozzáférési módok megkülönböztetésére, de a tartalmak létrehozására és etikai kérdésekre is.⁴ Az 1990-es évektől kezdve a klasszikus médiakörnyezet átalakult a hálózati kommunikáció hatására és egy új minőségű, konvergens médiakörnyezetté vált. Ezen átalakulás természetesen tükröződött a médiajártasság koncepcióiban is, hiszen az új, konvergens médiakörnyezet szükségszerűen új jártasságokat is magával hozott.⁵ Erre a változásra reflektált a médiakutatás, az akadémiai környezet és a pedagógia is, intézményi oldalról is jelentős támogatást kapva. Az Egyesült Nemzetek Nevelésügyi, Tudományos és Kulturális Szervezete, az UNESCO 2011-ben kifejezetten médiapedagógusok számára állított össze egy médiajártassággal foglalkozó összefoglaló művet, ahol a készségek körét Carolyn Wilson és munkatársai dolgozták ki.⁶ A kutatócsoport a tudás, az elemzőképeség és a létrehozás hármasságában gondolkodott, s ennek megfelelően dolgozták ki oktatási programjukat. A médiajártasságnak mint oktatási programnak a célja, hogy a diákok ismerjék és értsék meg a média és az információ fontosságát a demokratikus párbeszédben és a társadalmi folyamatok alakításában, az azokban való részvételben, bekapcsolódásban. Emellett képesek legyenek elemezni, feldolgozni, értékelni a médiaszövegeket, az információforrásokat. Amennyiben mé-

⁴ Szöke-Milinte Enikő: *Információ–Média (tudatosság)–Műveltség*. Budapest, Pázmány Péter Katolikus Egyetem, 2020.

⁵ Andok Mónika: Digitális média konvergencia – konvergens médiajártasságok. In: Hulyák-Tomesz Tímea (szerk.): *Generációs kérdések a kommunikációskészségfejlesztésben*. (A kommunikáció oktatása 11.) Budapest, Hungarovox, 2016. 41–49.; Andok Mónika: Kommunikációs kompetencia–médiajártasság. Koncepcionális különbségek, készségbeli hasonlóságok. In: H. Tomesz Tímea – H. Varga Gyula (szerk.): *Kommunikációs készségfejlesztés a pedagógiai munkában*. Budapest, Hungarovox, 2019. 31–39.

⁶ Az angol nyelvű szövegek fordításánál a szerző igénybe vett mesterséges intelligencia alkalmazásokat, kritikai ellenőrzés mellett. A szövegalkotásban nem kapott szerepet MI-alkalmazás.

diatartalmakat hoznak létre, azt felelősségteljes módon tegyék. Látható tehát, hogy a médiajártasság elemei a média működését társadalmi kontextusban ragadják meg, funkcionalitásában alapvetően normatív módon. A dokumentum a médiajártassághoz kapcsolódóan hat kompetenciakört azonosít, ezek a következők:

- (1) Megérteni a média és az információ szerepét a demokrácia működésében.
- (2) A médiatartalmak megértése és használata.
- (3) Az információkhoz való eredményes és hatékony hozzáférés képessége.
- (4) Az információk és az információforrások kritikus értékelésének képessége.
- (5) Az új és a hagyományos médiaformák ismerete, használata.
- (6) A médiatartalom szituációhoz, szociokulturális kontextushoz való kötöttségének kritikus ismerete.

A dokumentumból is leszűrhető, hogy a médiajártasság koncepciójába a 21.században már a hálózati kommunikáció, a közösségi média kérdései, az ehhez kapcsolódó jártasságok kapcsolódnak.⁷

1. ábra A médiajártasság elemei 2011-ben (forrás: Wilson et al. i. m.)

Figure 2: The Ecology of MIL: Notions of MIL



A 2010-es évek második felében született írások nagyobb egységekbe sorolják azon nélkülözhetetlen készségeket, melyek a hálózati, közösségi média felelős használatához

⁷ Carolyne Wilson. et. al.: *Media and Information Literacy. Curriculum for Teachers*. Paris, UNESCO, 2011.

elengedhetetlenek.⁸ Park megközelítése nyolc nagyobb készséget sorol fel, s a nyolc nagyobb jártasságkörön belül további három, kisebb hatókörű jártasságot is tartalmaz. Ezek a következők:

- (1) Digitális jogok: digitális szólásszabadság, szellemi termékekhez, alkotásokhoz fűződő jogok, a magánszféra védelme.
- (2) Digitális jártasság: a számítógépes logika és gondolkodás ismerete, a tartalom létrehozásának képessége és a kritikai gondolkodás.
- (3) Digitális kommunikáció: online együttműködés képessége, online kommunikációban való jártasság és digitális lábnyomok ismerete.
- (4) Digitális érzelmi intelligencia: a társas és érzelmi tudatosság fejlesztése, érzelmek kezelése és az empátia.
- (5) Digitális biztonság: a jelszavak védelme, a hálózati és a mobil használat biztonsága.
- (6) Digitális veszélyek elkerülése: a kapcsolati, a tartalmi és a viselkedési veszélyeket is beleértve.
- (7) Digitális használat: a képernyő előtt töltött idő kontrollja, a digitális egészség és a közösségi részvétel kérdése.
- (8) Digitális identitás: digitális alkotó, állampolgár és vállalkozó belső elemeit tudjuk elkülöníteni.

Park esetében már olyan elemeket is azonosíthatunk a médiajártasság összetevői között, melyek a közösségi média identitásalakító erejéhez kapcsolhatóak.

2. ábra: Nyolc konvergens digitális jártasság (forrás: Park i. m.)



⁸ Yuhyun Park: 8 digital skills we must teach our children. World Economic Forum, 2016. <https://tinyurl.com/chda78f3>

A technikai fejlődés, a mesterséges intelligencia⁹ előretörése szükségessé tette, hogy a médiajártasság mellett a mesterségesintelligencia-jártasságot is körültekintően áttekintsék a kutatók. A mesterséges intelligencia egy számítógépes rendszer azon képessége, hogy az emberihez hasonló kognitív funkciókat tud utánozni, amilyen például a tanulás vagy a problémamegoldás. Az MI koncepciója, első alkalmazásai az 1960-as években nem jelentettek azonnali áttörést az új technológia használatában.¹⁰ A mindennapokban tulajdonképpen a Chat GPT 2022-es megjelenése indította el a széles körű polgári alkalmazást. A mesterséges intelligencia korai definíciói az alapító atyák, Newell és Simon valamint Minsky nevéhez köthetőek. „Az »általános intelligens cselekvéssel« az intelligencia ugyanazon hatókörét kívánjuk jelezni, amelyet az emberi cselekvésben látunk: hogy minden valós helyzetben a céloknak megfelelően, a környezethez alkalmazkodva viselkedik a rendszer a komplexitás és a gyorsaság elérhető határain belül.”¹¹ Míg Minsky definíciója szerint az intelligencia általában a nehéz problémák megoldásának képességét jelenti.¹² A mesterséges intelligencia korai meghatározásait szélesítette ki 2019-ben Wang: szerinte a mesterséges intelligencia az, amely képes kognitív feladatokat ellátni, különösen tanulási és problémamegoldásbeli újításokkal él, mint pl. gépi tanulás, természetes nyelvi feldolgozás és neurális hálózatok.¹³ Fontos megemlíteni az AI-rendszerek azon sajátosságát, hogy idővel fejlődnek és alkalmazkodnak. Ennek eredményeként a mesterséges intelligencia tervezési ciklusa egy folyamatos újra tervezés, alakítás, ám ez folyamatos emberi felügyeletet is igényel, hogy ne sérüljenek az eredeti célok, s megmaradjon az összhang a felhasználói értékekkel és célokkal.¹⁴ A mesterségesintelligenciajártasság meghatározásánál természetesen figyelembe kell vennie az MI-hez kapcsolódó lehetőségeket és kihívásokat is. Ozlem Garibay és munkatársai hat olyan nagyobb területet említenek, ahol az MI erős kihívást jelent az emberi felhasználás tekintetében.¹⁵

⁹ A továbbiakban: gyakran MI, vagy az angol kifejezést (*artificial intelligence*) rövidítve AI.

¹⁰ Csepeli György: *Ember 2.0: A mesterséges intelligencia gazdasági és társadalmi hatásai*. Budapest, Kossuth, 2020.

¹¹ Allen Newell –Herbert A. Simon: Computer science as empirical inquiry: symbols and search. *Communications of the ACM*, Vol. 19., No. 3. (1976) 11–126.

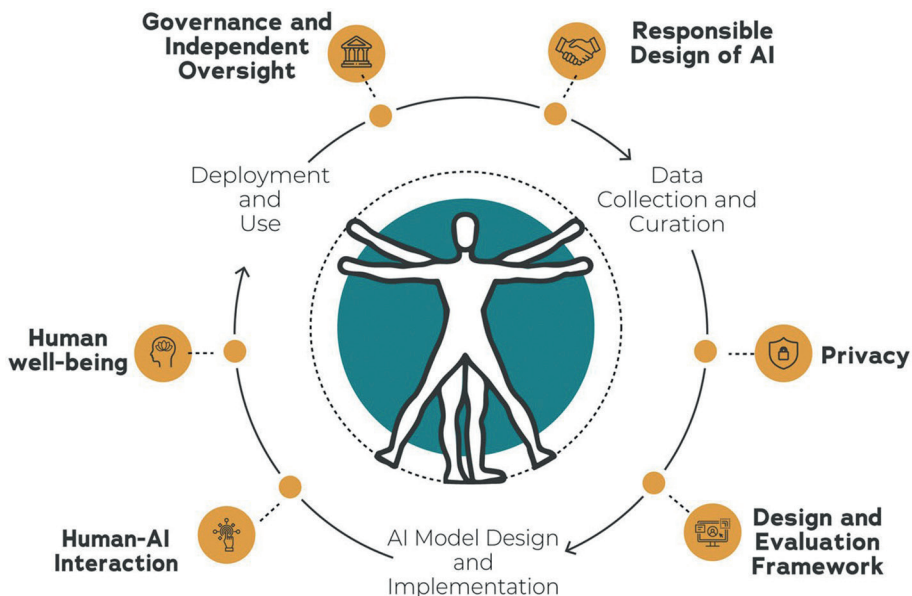
¹² Marvin Minsky: *Society of mind*. Simon and Schuster, 1988.

¹³ Pei Wang: On defining artificial intelligence. *Journal of Artificial General Intelligence*, Vol. 10., No. 2. (2019) 1–37.

¹⁴ Ozlem Ozmen Garibay – Brent Winslow – Salvatore Andolina – Margherita Antona – Anja Bodenschatz – Constantin Couraris et al.: Six Human-Centered Artificial Intelligence Grand Challenges. *International Journal of Human-Computer Interaction*, Vol. 39., No. 3. (2023) 391–437.

¹⁵ Uo.

3. ábra: A mesterséges intelligencia kihívásai (forrás: Ozmen Garibay et al. i. m. 393.)



Az emberközpontú mesterséges intelligencia rendszerek és technológiák kiépítésének hat nagy kihívása a mesterséges intelligencia fejlesztése, ami

- (1) emberi jóllét-orientált,
- (2) felelős,
- (3) tiszteletben tartja a magánéletet,
- (4) emberközpontú tervezést és értékelést tartalmaz,
- (5) lehetővé teszi az irányítást és a felügyeletet, és
- (6) tiszteletben tartja az emberi kognitív folyamatokat az ember-MI interakció hátterületén.

Az első nagy kihívás azt célozza, hogy a mesterséges intelligencia rendszerek az emberek életének és tapasztalatainak javítására összpontosítsanak. A második és harmadik kihívásterület olyan alapelveket képvisel, amelyek biztosítják a felelős MI-rendszertervezést és -fejlesztést, valamint védik az emberi magánéletet. A negyedik és ötödik kihívás a mesterséges intelligencia életciklusa feletti megfelelő útmutatást és emberi ellenőrzést szolgáló átfogó keretrendszer kidolgozását és biztosítását jelenti. Az utolsó kihívás pedig a jövőbeli human-AI interakció víziója. De természetesen ezek a kihívások nem szegmentálhatók ilyen pontossággal, sokkal inkább jellemző rájuk az átfedés, az összekapcsolódás.

3. Mesterségesintelligencia-jártasság a tanulók, hallgatók szempontjából

A médiajártassághoz hasonlóan a mesterséges intelligencia használatának gyakorlottságára vonatkozóan is megjelentek konceptualizációk a szabályozás és a médiakutatás irányából is. Az MI-jártasság jelentőségét az is alátámasztja, hogy az EU mesterséges

intelligenciáról szóló rendelete 3. cikk Fogalommeghatározások 56. pontja is definiálja, mégpedig a következőképpen:

„56. »MI-jártasság«: olyan készségek, ismeretek és értelmezési képességek, amelyek lehetővé teszik a szolgáltatók, az alkalmazók és az érintett személyek számára, hogy – figyelembe véve az e rendelettel összefüggésben fennálló jogait és kötelezettségeiket – a megfelelő információk birtokában telepítsenek MI-rendszereket, valamint tudomást szerezzenek az MI lehetőségeiről és kockázatairól, és azon lehetséges károkról, amelyet az MI okozhat;”¹⁶

A MI-jártassághoz kapcsolódó, annak ismeretére fókuszáló, pedagógiai jellegű első szisztematikus áttekintést 2021-ben Ng és munkatársai írták le, akik 30 olyan kutatást vizsgáltak meg, melyek az MI-jártasság konceptualizálását célozták.¹⁷ Négy területet emelnek ki a szakirodalmak és kutatások alapján: ismerje és értse a mesterséges intelligenciát, használja és alkalmazza a mesterséges intelligenciát, kritikusan értékelje és hozzon létre MI-t, s legyen tisztában az MI etikus használatával. Az első pont azt tartalmazza, hogy a hallgatókat meg kell tanítani az MI használatára. Oktatási feladatként lett tehát megfogalmazva a mesterséges intelligenciához kapcsolódó alapvető fogalmak, készségek, ismeretek és attitűdök elsajátítása. Bár a hallgatók az MI végfelhasználói, meg kell érteniük a mögöttes technológiákat. A második kompetencia-kör az MI használatát fogja össze, s olyan elvárásokat támaszt, hogy a hallgató tudja, hogyan alkalmazza a mesterséges intelligenciát különböző helyzetekben, környezetekben, akár az oktatásban, akár a mindennapi életben. Ugyanakkor képesnek kell lenniük arra, hogy kritikusan értékeljék az MI-technológiákat, illetve képesek legyenek hatékonyan és etikus módon együttműködni az MI alkalmazásokkal. Kutatásukban két korábbi keretrendszert is bemutatnak, az egyik Bloom – eredetileg az 1950-es években kidolgozott – taxonómiája az MI-jártasságra adaptálva, a másik pedig a TPACK keretrendszer.¹⁸

¹⁶ Az Európai Parlament és a Tanács (EU) 2024/1689 rendelete (2024. június 13.) a mesterséges intelligenciára vonatkozó harmonizált szabályok megállapításáról, valamint a 300/2008/EK, a 167/2013/EU, a 168/2013/EU, az (EU) 2018/858, az (EU) 2018/1139 és az (EU) 2019/2144 rendelet, továbbá a 2014/90/EU, az (EU) 2016/797 és az (EU) 2020/1828 irányelv módosításáról (a mesterséges intelligenciáról szóló rendelet) (EGT-vonatkozású szöveg). HL L, 2024/1689, 2024.7.12., 3. cikk 56. pont. <https://tinyurl.com/2ynnxbm6>

¹⁷ Davy Tsz Kit Ng – Jac Ka Lok Leung – Samuel Kai Wah Chu – Maggie Shen Qiao: Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence*, Vol. 2. (2021) 100041. <https://doi.org/10.1016/j.caeai.2021.10004>

¹⁸ Ng et al. i. m. 2.

4. ábra: Bloom mesterséges intelligencia jártassági ábrája [forrás: Ng et al. i. m. 5.]

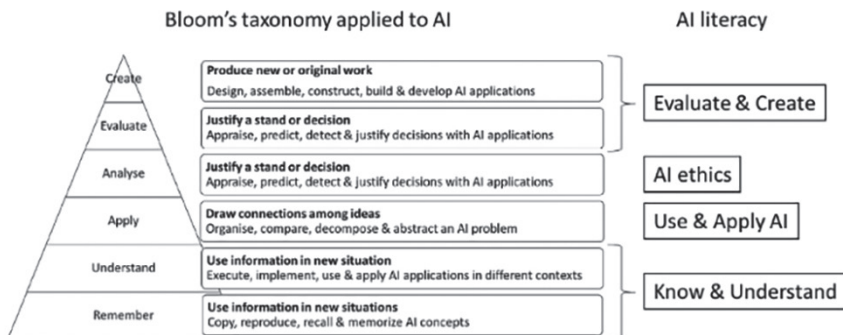
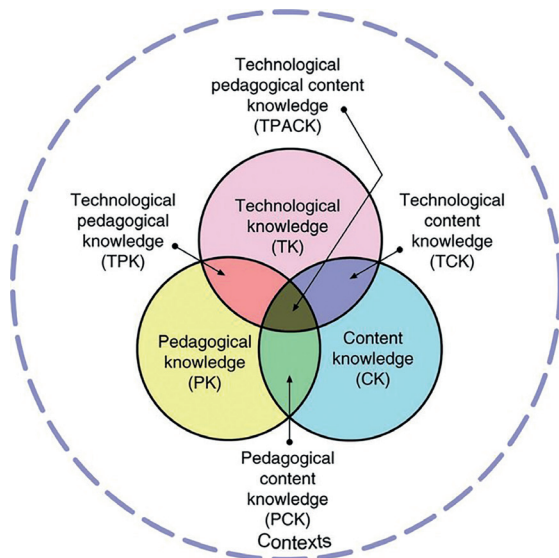


Fig. 1. Bloom's Taxonomy and AI literacy adapted from Ng et al. (2021).

5. ábra: TPACK (Technological pedagogical content knowledge) mesterségesintelligencia-jártassági ábrája [forrás: Ng et al. i. m. 5.]



Az eddig ismertetett kompetenciakörök alapján fontos hangsúlyoznunk, hogy az MI-hez kapcsolódóak hasonlítanak a klasszikus digitális kompetenciánál már felsoroltakra; lényegében a megértés, a használat és az létrehozás folyamatában gondolkodnak. Változás viszont a korábbi, digitális médiakompetenciákhoz képest az etika erőteljesebb megjelenítése a kompetenciakörökben. S látni fogjuk, hogy a következőkben ismertetendő koncepciókban is az etikai kérdések fókuszáltságát fogjuk találni. Jelen tanulmány legrészletesebben az UNESCO által kidolgozott MI-jártasságra vonatkozó

keretrendszer mutatja be és elemzi.¹⁹ Az UNESCO a médiajártasság keretrendszerének kidolgozásához hasonlóan a mesterséges intelligencia kapcsán is elkészítette a javaslatát 2024-ben, mind a pedagógusok, felsőoktatásban dolgozók, mind a diákok, a hallgatók számára.

Az UNESCO-dokumentum MI-jártasság szükségességéről szóló érvelésében azt olvashatjuk, hogy mivel a mesterséges intelligencia egyre inkább szerves részét képezi életünknek, ezért proaktív oktatási rendszerekre van szükség ahhoz, hogy felkészítsék a tanulókat az MI felelős felhasználására és alakítására. Az általuk kidolgozott keret elnevezése „*AI competency framework for students*”, rövidítése pedig AI CFS. Megjegyzik azt is, hogy 2022-ig csak 15 ország építette be a mesterséges intelligencia tanulási célkitűzéseit nemzeti tantervébe. A MI-kompetencia három fő pilléren nyugszik, amelyek szélesebb alapkompenciákat határoznak meg a diákok számára, nevezetesen a tudás, a készség és az értékek.²⁰

Az UNESCO tanulói MI-kompetenciakerete 12 kompetenciakört vázol fel, négy területen, három szinten elkülönítve. A négy terület a következő: az (1) emberközpontú gondolkodásmód, (2) etikus MI, (3) MI-technikák és -alkalmazások, valamint az (4) MI rendszertervezés. Ezek a kompetenciák három haladási szinten jelenhetnek meg: a megértés, az alkalmazás és a létrehozás szintjén. Ugyanakkor mivel az MI folyamatosan változik, alakul, fejlődik, ezt a kompetenciák leírásában is meg kell jeleníteni. A keretrendszert kidolgozók szándéka szerint az AI CFS egyrészt globális referenciakeretet biztosít az MI-kompetenciák megjelenítéséhez a hallgatók számára. Másrészt meghatározza a MI-kompetenciák kulcsfontosságú szempontjaihoz kapcsolódó tipikus attitűdöt és viselkedést a különböző elsajátítási szinteken, hogy segítse az MI-vel kapcsolatos tantervi anyag megtervezését a hallgatók számára. Harmadrészt pedig nyitott ütemtervet ajánl az MI-tantervek tanulási sorrendjének megtervezéséhez az egyes évfolyamokon.²¹

Az első területként definiált emberközpontú gondolkodásmód azt jelenti, hogy az MI használatának végső célja nem a profit termelés, nem gazdasági célt kell látni, hanem az ember szolgálatát. Az UNESCO emberközpontú megközelítése azt hirdeti, hogy a mesterséges intelligencia tervezésének és használatának az emberi képességek fejlesztését, a lehetőségek szélesre nyitását kell szolgálnia. Emellett meg kell védenie az emberi méltóságot, az önrendelkezést, valamint elő kell mozdítania az igazságosságot és a fenntarthatóságot a mesterséges intelligencia teljes életciklusa során. Ezt a megközelítést az emberi jogok, valamint a közös tudást meghatározó nyelvi és kulturális sokszínűség tiszteletben tartásának kell vezérelnie – fogalmaz a dokumentum. Az emberközpontú megközelítés gondolatában az is benne foglaltatik, hogy a mesterséges intelligenciát átlátható módon, az emberi ellenőrzés lehetőségét fenntartva kell használni.²² A második szempont a mesterséges intelligencia etikája, mely magában

¹⁹ AI competency framework for students. UNESCO 2024. <https://doi.org/10.54675/JKJB9835> [a továbbiakban: UNESCO (2024a)].

²⁰ Uo. 18.

²¹ Uo.14.

²² Uo.15.

foglalja a tanulók MI-kompetenciájának társas és etikai összetevőit, ideértve az alkalmazásához való emberi, felhasználói hozzájárulás szükségességét. A harmadik szempont a mesterséges intelligencia alkalmazásához kapcsolódó készségeket, s az egymással szorosan összefüggő fogalmi ismereteket jelenti. Pontosan azt, hogy tudatosan kiválasztott MI-eszközök segítségével oldanak meg feladatokat a hallgatók. Az utolsó szempont pedig az MI rendszertervezés, amely átfogó mérnöki készségeket takar, amelyek meghatározzák a megoldandó problémát, a technológiai architektúra felépítését, a képzést, a tesztelést, valamint az MI-rendszerek optimalizálását.

A keretrendszer második dimenziója az előrehaladás három szintjét vázolja fel, melyek a megértés, az alkalmazás és a létrehozás szintjei.

1. táblázat: AI kompetenciakeret a tanulók számára [forrás: UNESCO (2024a) i. m. 19.]

Kompetencia szempontok	Haladási szintek		
	Megértés	Alkalmazás	Alkotás
Emberközpontú gondolkodásmód	Emberi ágencia	Emberi elszámoltathatóság	Állampolgárság a mesterséges intelligencia korszakában
A mesterséges intelligencia etikája	Etikai kérdések átlátása az MI tekintetében	Biztonságos és felelősségteljes használat	Tervezési etika
AI technikák és alkalmazások	MI alapok	Alkalmazási ismeretek	MI-eszközök létrehozása
AI rendszertervezés	Probléma hatóköre	Architektúra tervezés	Iterációs és visszacsatoló hurkok

3.1. Emberközpontú gondolkodás

Az emberközpontú gondolkodás három szinten, a következőképpen jelenik meg az UNESCO anyagában: a megértés szintjén az emberi ágencia kérdését érintik, az alkalmazás szintjén az elszámoltathatóságot és az alkotás szintjén az állampolgársághoz kapcsolódóan megjelenő MI-tervezést. Az emberi ágencia szerint a tanulóktól elvárják, hogy felismerjék: a mesterséges intelligencia ember által vezérelt, és hogy az MI-alkotók döntései befolyásolják azt a módot, ahogyan az MI-rendszerek befolyásolják az emberi jogokat, az ember és a mesterséges intelligencia interakcióját, valamint saját életüket, közösségeiket és társadalmukat. Ezzel együtt a hallgatóknak azt is fel kell tudni mérniük, hogy milyen következményekkel, veszélyekkel járhat, ha nem emberközpontú az MI használata, alkalmazása. Az alkalmazás szintjén az emberi elszámoltathatóság alatt azt értik, hogy a hallgatóknak fel kell ismerniük, hogy az emberi felelősségvállalás a mesterséges intelligencia alkotóinak és szolgáltatóinak jogi kötelezettsége. Egyúttal felelősséget kell vállalniuk az MI tervezése és használata során. A digitális felületeken is felelősen viselkedő állampolgári felfogáshoz kapcsolódóan a mesterséges intelligencia korszakában a tanulóktól elvárják, hogy kritikusan értsék meg a mesterséges intelligencia emberi társadalmakra gyakorolt hatását, és támogassák a mesterséges intelligencia felelős, befogadó tervezését és használatát a fenntartha-

tó fejlődés érdekében. Cél az is, hogy a hallgatók – az élethosszig tartó tanulás keretében – folytassák a mesterséges intelligencia megismerését és használatát.²³

3.2. Az MI etikája

A dokumentum második nagy kompetenciaköre az etikához kapcsolódik. Az első szinten található az MI etikai vonatkozású kérdéseinek átlátása. Vagyis a hallgatóktól elvárják, hogy alapvető ismereteket szerezzenek a mesterséges intelligenciával kapcsolatos kulcsfontosságú etikai viták alapjául szolgáló kérdésekről. Beleértve a mesterséges intelligenciának az emberi jogokra, a társadalmi igazságosságra, a befogadásra, a méltányosságra és az éghajlatváltozásra gyakorolt hatását mind környezetükben, mind személyes életükben. Ehhez a szinthez köti még a dokumentum, hogy a hallgatók ne csak megértsék, de internalizálják és reflektív módon alkalmazzák is a következő alapelveket: a káros, ártó használat elkerülése, az arányosság és a diszkrimináció kerülésének elve, illetve a fenntarthatóság elve. Még pontosabban körvonalazva az ártalmas használat elkerülése azt jelenti, hogy a hallgatók megértik, hogy az MI-rendszereket nem szabad olyan célokra használni, amelyeknek emberekre nézve káros következményeik lehetnek (például arcfelismerés, társadalmi státusz hozzárendelése, vagy prediktív algoritmusok osztályozási vizsgákhoz). Az arányosság etikai elve azt jelenti, hogy a hallgatók a saját igényeiknek megfelelően képesek annak vizsgálatára, hogy egy adott MI-rendszer alkalmazása előnyös-e, indokolt-e egy cél eléréséhez, illetve, hogy egy adott AI-módszer megfelelő-e az adott kontextusban. A diszkrimináció kerülésének elve alapján a hallgatók tisztában vannak a nemi, etnikai, kulturális és egyéb torzításokkal, amelyek az MI-eszközökbe vagy azok kimeneteibe ágyazódnak, és képesek felismerni azokat. A fenntarthatóság etikai elve alapján a tanulók képesek elmagyarázni és szemléltetni az MI-rendszerek környezeti fenntarthatóságra gyakorolt hatásait.²⁴ Az ember és a mesterséges intelligencia együttműködéséhez kapcsolódóan a tanulónak meg kell érteniük, miért kell az embereknek etikai és jogi felelősséget vállalniuk az MI használatáért; s hogyan lehetnek elszámoltathatóak a mesterséges intelligencia által támogatott döntéshozatali helyzetekben.²⁵ Mindezek mellett a dokumentum még ehhez a szinthez köti a MI-jártasság átláthatóságának és megmagyarázhatóságának kérdését is.

Az MI etikájához kapcsolódó haladó készségként határozza meg a dokumentum a biztonságos és felelősségteljes használatot. Ezek alapján a hallgatóktól elvárható, hogy az etikai alapelveknek és előírásoknak megfelelően, felelősségteljesen tudják használni az MI-t. Vagyis tisztában vannak az adatvédelem kérdéseivel éppúgy, mint az adatok feltárásával járó kockázatokkal. Tisztában vannak bizonyos MI-rendszerek sajátos kockázataival is, és képesek megvédeni saját, valamint társaik biztonságát az MI használata során. Az MI etikájának legmagasabb szintjén az alkotást, az etikus tervezést találjuk. A tervezési etika kérdéskörébe tartozik, hogy a hallgatók felé elvárás, hogy

²³ Uo. 22.

²⁴ Uo. 24.

²⁵ Uo. 25.

az AI-eszközök tervezése, értékelése és használata, valamint a mesterséges intelligencia szabályozásának felülvizsgálata és adaptálása során az etika beépítésén alapuló megközelítést alkalmazzanak. Ezzel együtt a tanulóknak tisztában kell lenniük azzal is, hogy a mesterséges intelligencia tervezése mögött meghúzódó szándék felmérése magában foglalja az AI teljes életciklusát, a koncepcióalkotás szakaszától kezdve. A tanulóknak képesnek kell lenniük felmérni, hogy egy AI-eszköz megfelel-e az etikai előírásoknak, valamint át kell tekinteni a mesterséges intelligencia szabályozásait, és tájékoztatni kell az alkalmazókat a kockázatokról

3.3. Technikák és alkalmazások

Az UNESCO keret harmadik nagy kompetenciaköre az MI technikákat és alkalmazásokat öleli fel, a mesterséges intelligencia működésének egymással szorosan összefüggő fogalmi ismereteit jelenti.²⁶ Az alapok szintjén ez azt jelenti, hogy a hallgatóktól elvárják, hogy alapvető ismeretekre és készségekre tegyenek szert az MI-vel kapcsolatban, különösen az adatok és algoritmusok tekintetében. Ezzel együtt a hallgatóknak képesnek kell lenniük a mesterséges intelligencia fogalmi ismereteinek összekapcsolására a társadalmi folyamatokkal és a mindennapi életben végzett tevékenységeikkel, konkretizálva az emberközpontú gondolkodásmódot és az etikai elveket azáltal, hogy megértik, hogyan működik a mesterséges intelligencia, hogyan kommunikál, illetve kerül interakcióba az emberekkel. Középszinten (haladó) már az alkalmazási készségekkel számol az AI CFS keretrendszer. Ezen a szinten már elvárják a tanulóktól, hogy képesek legyenek az adatok, az MI-algoritmusok és a programozás megértésére, hogy kritikusan értékeljék és használják az ingyenes és/vagy nyílt forráskódú MI-eszközöket, programkönyvtárakat és adatbázisokat.

Az alkotás szintjén jelenik meg az MI-eszközök létrehozásának készsége. Itt a hallgatóktól már elvárják, hogy ne csak elmélyítsék és alkalmazzák az adatokkal és algoritmusokkal kapcsolatos ismereteiket és készségeiket, hanem képesek legyenek testre szabni a meglévő MI-eszközöket, sőt képesek legyenek feladat alapú MI-eszközök létrehozására.

A négy nagy kompetenciakör összekapcsolódását a dokumentumban az is jelzi, hogy a hallgatóktól elvárják, hogy integrálják az emberközpontú gondolkodásmódot és etikai szempontjaikat a meglévő MI-források értékelésébe. Azt is elvárják tőlük, hogy fejlesszék azokat a társas és érzelmi készségeket, amelyek az MI-vel való alkotáshoz szükségesek, beleértve az alkalmazkodóképességet, a komplex kommunikációs és csapatmunka készségeket.

3.4. Rendszertervezés

Az AI CFS negyedik kompetenciaköre az MI rendszerek tervezésének aspektusa a rendszerszintű tervezésre és az MI-rendszerek optimalizálásához szükséges átfogó

²⁶ Uo.

mérnöki készségekre összpontosít.²⁷ Itt az első szinten a megértést, a probléma hatókörének meghatározását várja el a dokumentum. A hallgatóknak képesnek kell lenniük annak megértésére, hogy mi adhatja az MI-innováció kiindulópontját konkrét fejlesztések esetében. Elvárják, hogy jogi, etikai és logikai szempontból megvizsgálják, hogy bizonyos helyzetekben érdemes-e a mesterséges intelligenciát használni, mielőtt megpróbálna egy MI-modellt betanítani a megoldására. A hallgatóktól elvárják továbbá, hogy elsajátítsák a MI-rendszer koncepciójának kialakításához és felépítéséhez szükséges ismereteket és projekttervezési készségeket, beleértve a különböző MI-technikák megfelelőségének felmérését, az adatigények meghatározását, valamint a teszt- és vizsacsatolási mérőszámok kidolgozását. Középszintű készségként fogalmazódik meg az AI CFS keretrendszerében az architektúra tervezés. Ezek alapján a hallgatóknak képesnek kell lenniük az alapvető módszertani ismeretek és technikai készségek fejlesztésére egy skálázható, karbantartható és újra felhasználható architektúra konfigurálásához. A hallgatóktól elvárják, hogy fejlesszék azokat az interdiszciplináris készségeiket, amelyek az adatkészletek, programozási eszközök és számítási erőforrások kihasználásához szükségesek egy MI-rendszer megalkotásához. Az alkotás szintjén a hallgatóknak képesnek kell lenniük életkoruknak megfelelő technikai készségek elsajátítására az adatbázisok fejlesztéséhez, az algoritmusok újra konfigurálásához és az architektúrák fejlesztéséhez a tesztek eredményeire és a visszajelzésekre reagálva. Az AI CFS a kompetenciakörök megfogalmazásán, kidolgozásán túl konkrét tantervi célkitűzéseket és pedagógusi módszereket is javasol a kompetenciák kialakításához, mind alap-, mind haladó, illetve az alkotás szintjén.²⁸ S végezetül szakterület-specifikus értékeléseket is javasol a szöveg a három előrehaladási szinten.²⁹

4. MI kompetenciák a pedagógusok számára

A MI-jártasság kérdése nemcsak a hallgatók, de az oktatók tekintetében is a figyelem fókuszába került a 2020-as évek elejétől, s megkezdődött az MI-műveltség (*AI literacy*) pedagógusokra adaptált változatainak a kidolgozása. Ezek célja, hogy az MI hatékonyan integrálható legyen a tanulási-tanítási folyamatokba, s a pedagógusok szakmai fejlődésének támogatnia kell az MI technológia, pedagógiai és szaktárgyi értelmezését.³⁰ Lamerás és munkatársai taxonómiájában a pedagógusokra vonatkozó kompetenciák a következők:

- digitális tartalom tervezése, fejlesztése és szolgáltatása;
- adatszerzés, információszerzési és adatetikai készségek megszerzése;
- digitális tevékenység-vezérelt pedagógia alkalmazásával kapcsolatos készségek fejlesztése;
- az oktatásban használható MI-alkalmazások, -eszközök és -szoftverek használatának elsajátítása;

²⁷ Uo. 26.

²⁸ Uo. alapszinten 29–36., haladó szinten 37–44., az alkotás szintjén 45–52.

²⁹ Uo. 72.

³⁰ Ng et al. i. m.

- kreatív digitális készségek, fejlesztése; a digitális inklúzió erősítése;
- a diákok társadalmi felelősségvállalásának és tudatos adatkezelésének elősegítése.³¹

2024-ben az UNESCO a hallgatókra vonatkozó MI kompetencia keretrendszerrel párhuzamosan a pedagógusokra vonatkozó elvárásokat is kidolgozta, *AI CFT Competency framework for teacher* néven.³² Célja, hogy segítse a tanárokat megérteni az ember és mesterséges intelligencia interakcióval kapcsolatos többretegű és több dimenziójú értékekkel és attitűdök kapcsolatban.³³

A tanárok MI-kompetenciájának keretrendszere a hallgatókéhoz hasonlóan két-dimenziós, az egyik dimenzióban a kompetencia öt aspektusát fogalmazzák meg, a másik dimenzióban pedig mindezt három előrehaladási szinten, így összességében tizenöt blokkot alkotva, amint az 2. táblázatban látható.

Az első dimenzió az MI-kompetencia öt aspektusát tartalmazza, amelyek a bal szélen láthatók a 2. táblázat első oszlopában. A kompetenciaszempontok a tudás, készségek, értékek és attitűdök egymással összefüggő kulcselemeit jelentik, amelyeket a tanároknak ki kell fejleszteniük ahhoz, hogy a mesterséges intelligenciát hatékonyan és etikusan integrálhassák tanítási gyakorlatukba, mind a tanulás elősegítése mind a szakmai fejlődésük terén. Ez az öt szempont: az emberközpontú gondolkodásmód, a mesterséges intelligencia etikája, a mesterséges intelligencia alapjai és alkalmazásai, a mesterséges intelligencia pedagógiája és a szakmai fejlődést szolgáló mesterséges intelligencia. Vagyis az öt kompetenciakörből az első három megegyezik a hallgatók számára megfogalmazott kompetenciakörökkel, a negyedik és az ötödik vonatkozik csak a pedagógusokra. Az első aspektus, az emberközpontú gondolkodásmód, meghatározza azokat az értékeket és attitűdöket, amelyeket a tanároknak követniük kell. A második aspektus a mesterséges intelligencia etikája, ez tartalmazza azokat az alapvető etikai elveket, intézményi előírásokat és gyakorlatokat, amelyeket a tanároknak meg kell érteniük, alkalmazniuk és alkalmaztatniuk kell. A harmadik a fogalmi ismereteket és az átadható készségeket határozza meg, amelyeket a tanároknak meg kell érteniük és alkalmazniuk kell az MI-eszközök kiválasztása, az alkalmazása során. A negyedik aspektusa az AI CFT-nek az MI célirányos és hatékony pedagógiai integrációjához szükséges képességeket írja le. Ez magában foglalja a megfelelő MI-eszközök kiválasztásának képességét, valamint a pedagógiai stratégiákba való integrálását a kurzusok előkészítésénél, kivitelezésénél és a hallgatói értékelésnél. S végezetül, ötödik aspektusként említik a szakmai fejlődést, s felvázolják azokat a kompetenciákat, amelyeket a tanároknak fejleszteniük kell ahhoz, hogy megfelelően használják az AI-t élethosszig tartó szakmai fejlődésük során.

³¹ Petros Lamerás – Sylvester Arnab: Power to the teachers: an exploratory review on artificial intelligence in education. *Information*, Vol. 13., No. 1. (2021) 14.

³² AI competency framework for teachers. UNESCO, 2024. <https://doi.org/10.54675/ZJTE2084> [a továbbiakban: UNESCO (2024b)]

³³ Uo. 17.

2. táblázat: Az MI kompetencia keretrendszer magas szintű felépítése: szempontok és előrehaladási szintek [forrás: UNESCO (2024b) i. m. 22.]

Szempontok	Haladás		
	Elsajátítás	Elmélyítés	Létrehozás
1. Emberközpontú gondolkodásmód	Emberi ágencia	Az emberi felelősség	Társadalmi felelősségvállalás
2. Az MI etikája	Etikai alapelvek	Biztonságos és felelősségteljes használat	Etikai szabályok közös megalkotása
3. MI alapok és alkalmazások	Alapvető MI technikák és alkalmazások	Alkalmazási készségek	Létrehozás MI-vel
4. MI pedagógia	MI által támogatott oktatás	MI – pedagógia integráció	MI által továbbfejlesztett pedagógiai átalakítás
5. MI a szakmai fejlődéshez	Az egész életen át tartó professzionális tanulást lehetővé tevő mesterséges intelligencia	MI a szervezeti tanulás fokozására	MI a szakmai átalakulás támogatására

2025-ben öt tudományos adatbázis (*ACM Digital Library, IEEEExplore, Clarivate Web of Science, Scopus, Dimensions*) MI-jártasságra vonatkozó tanulmányait dolgozta fel munkatársaival Annapureddy.³⁴ Eredményeik nyomán egy 12 elemű MI-jártassági keretet azonosítottak. Míg az UNESCO által javasolt keret elméleti konstrukcióként fogalmazódott meg, addig Annapureddy-éké empirikus kutatások összesítésén alapszik. Emiatt meglátásom szerint az utóbbi „életszerűbb”, felhasználó-centrikusabb megközelítést tartalmaz.

3. táblázat: MI-jártasság keretrendszere Annapureddy alapján (forrás: Annapureddy et al. i. m. fig. 1.)

Folyamatos tanulási, fejlődési képesség az MI alkalmazása során	Alapvető MI-műveltség		
	Generatív MI modellek ismerete	A generatív MI-eszközök kapacitásának és korlátainak ismerete	A generatív MI-eszközök használatának készsége
	Képes az MI által generált tartalom észlelésére		
	Képesség a generatív MI-eszközök kimenetének, létrehozott tartalmainak értékelésére	Generatív mesterséges intelligencia-eszközök használatának készsége (prompt engineering)	Programozási és finomhangolási képesség
A generatív MI használat kontextusainak ismerete	Az etikai vonatkozások ismerete	Jogi szempontok ismerete	

³⁴ Ravinithesh Annapureddy – Alessandro Fornaroli – Daniel Gatica-Perez: Generative AI literacy: Twelve defining competencies. *Digital Government: Research and Practice*, Vol. 6., No. 1. (2025) 1–21.

Az Annapureddy-ék által kidolgozott modell, MI-jártassági keretrendszer jó kiindulópontként szolgálhat a felsőoktatás számára, hogy az oktatók felmérjék majd akár diszciplína-specifikusan bővítsék a hallgatók MI alkalmazási készségeit.

5. Az AI-műveltség mérése

A tanulmány következő fejezésében az MI-jártasság méréséhez rendelkezésre álló eszközöket mutatjuk be. Az elmúlt öt évben a kutatók különféle skálákat fejlesztettek ki az MI-műveltség felmérésére.³⁵ Tekintettel arra, hogy az MI-műveltség kutatása túlsúlyban van az oktatási kontextusban, ezért sok mérőeszközt eleve oktatási környezetre terveztek. Ezekben a mérési eszközökben gyakori, sőt dominál az egyválasztásos vagy nyílt végű tudástesztek alkalmazása.³⁶ Ám az ilyen értékelések korlátja, hogy hajlamosak szorosan igazodni az értékelendő tananyag tartalmához. Alternatív megoldásként egyes kutatók önértékelési eszközöket alkalmaznak, míg mások a tudástesztek és önértékelések kombinációját javasolják, mindkét megközelítés erősségeinek kiaknázására.³⁷ Megjelentek differenciált tesztek kifejezetten a közoktatásban tanulók, a felsőoktatásban tanulók és a teljes lakosság körében való alkalmazásra is. Ezek közül néhányat említve, az AI-CI – *AI Literacy Concept Inventory Assessment*, AILQ – *AI Literacy Questionnaire*, AILS – *AI Literacy Scale*, AISES – *AI Self-Efficacy Scale*. Az első, az AI-CI egy teljesítményalapú koncepció-leltár, amelyet arra terveztek, hogy feleletválasztós kérdések segítségével értékelje a középiskolai diákok mesterséges intelligenciával kapcsolatos tudását. A skála tartalmát a szerzők MI-műveltsége alapján alakították ki. Az AI-CI lényegesen általánosabb és kevésbé kontextusfüggő az AI-műveltségi teszthez képest.³⁸ A másodikként említett AILQ középiskolai tanulók számára készült, és magában foglalja az AI-műveltség kognitív területeit, amelyek a szerzők feltáró áttekintéséből származnak, az érzelmi, viselkedési és etikai tanulási területek mellett.³⁹ A harmadik, az AILS az általános lakosságot célozza meg az ember-AI interakció (HAI) összefüggésében. A skála keretrendszere a MI-műveltségre vonatkozó koncepción alapul, amely szakirodalmi áttekintésből származik, s négy elemet azonosít: a tudatosságot, a használatot, az értékelést és az etikát.⁴⁰ AISES a mesterséges intelligencia önhatékonyságát méri fel az általános populációban.

³⁵ Tomás Lintner: A systematic review of AI literacy scales. *npj Science of Learning*, Vol. 9., No. 1. (2024) 50.

³⁶ Bingcheng Wang – Pei-Luen Patrick Rau – Tianyi Yuan: Measuring user competence in using artificial intelligence: validity and reliability of artificial intelligence literacy scale. *Behaviour & information technology*, Vol. 42., No. 9. (2023) 1324–1337.

³⁷ Duri Long– Brian Magerko: What is AI literacy? Competencies and design considerations. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 2020. 1–16. <https://doi.org/10.1145/3313831.3376727>

³⁸ Marie Hornberger – Arne Bewersdorff – Claudia Nerdel: What do university students know about Artificial Intelligence? Development and validation of an AI literacy test. *Computers and Education: Artificial Intelligence*, Vol. 5. (2023) 100165.

³⁹ Ng et al. i. m.

⁴⁰ Celalettin Çelebi – Faith Yılmaz – Uğur Demir – Ferhat Karakuş: Artificial intelligence literacy: An adaptation study. *Instructional Technology and Lifelong Learning*, Vol. 4., No. 2. (2023). 291-306.

Kialakulását a négytényezős szerkezetű, technológiával összefüggő önhatékonyság korábbi kutatásai alapozzák meg.⁴¹ A 2024-ig kidolgozott tesztekéről Lintner ad átfogó képet, felsorolja és röviden ismerteti a jelentősebbeket.⁴² Jelen fejezetben a két legelterjedtebb mérési eszközt mutatom be, az egyik az AILQ teszt, mely Ng és munkatársai nevéhez köthető, a másik a MALIS teszt, melyet Carolus és munkatársai dolgoztak ki.⁴³

Az AILQ teszt, ami a MI-műveltség-mérő eszköz, Ng és munkatársai fejlesztették ki 2023-ban, s a tanulók műveltség-fejlődésének négy dimenzióban történő mérésére.⁴⁴ Ng és munkatársai az MI-műveltségi kérdőívek szakirodalmának áttekintése után az azonosított kompetenciákat négy dimenzióba sorolták, melyek a következők: az affektív tanulás, a viselkedési tanulás, a kognitív tanulás és az etikus tanulás. Ezt egy 32 elemből álló, önbevallásos kérdőív segítségével tesztelték, mérték.

A másik bemutatandó mérőeszköz a Carolus és munkatársai által kifejlesztett teszt, a MAIIS (*Meta AI Literacy Scale: AI Literacy Questionnaire*) jól megalapozott kompetenciamodelleken és pszichológiai változás- és meta-kompetenciákon alapul.⁴⁵ A korábban alkalmazott 8 faktor helyett, melyből 4 a szakirodalomból levezett és 4 konkrét pszichológiai szempont volt, ők két magasabb funkciójú tényezőre, az „MI-műveltségre” és az „MI-önmenedzsmentre” bontották a koncepciót. Az MI-műveltség alá négy elemet soroltak: a használat és alkalmazás; a tudás, megértés; a kiértékelés, létrehozás és az MI-etika kérdéseit. (Vegyük észre, hogy ezekben nagy mértékben viszsza-köszön az UNESCO AI CFS keretrendszere.) Az MI-önmenedzsment alá sorolták az MI segítette problémamegoldást, az MI-tanulást, az MI segítette meggyőzést és az érzelemszabályozást. Ezzel a koncepcióval, saját értékelésük szerint is, nagymértékben eltérnek Ng és munkatársai koncepciójától.

6. ábra: MI-kompetencia területei és mérése Carolus és munkatársai alapján
(Forrás: Carolus et al. i. m.)

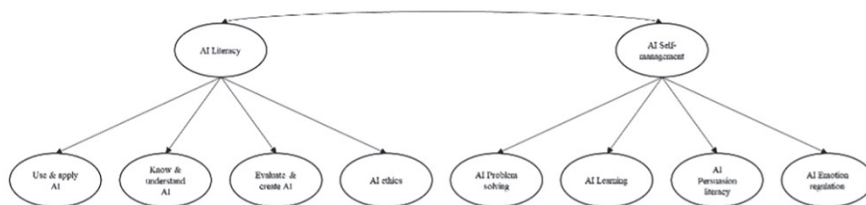


Fig. 1. Conceptual model for the confirmatory factor analysis.

⁴¹ S. Cassidy – P. Eachus: Developing the computer user self-efficacy (CUSE) scale: Investigating the relationship between computer self-efficacy, gender and experience with computers. *Journal of Educational Computing Research*, Vol. 26., No. 2. (2002) 133–153.

⁴² Lintner i. m.

⁴³ Astrid Carolus – Martin Koch – Samantha Straka – Marc Erich Latoschik – Carolin Wienrich: MAIIS- Meta AI literacy scale: Development and testing of an AI literacy questionnaire based on well-founded competency models and psychological change-and meta-competencies. *Computers in Human Behavior: Artificial Humans*, Vol. 1., No. 2. (2023) 100014. <https://doi.org/10.48550/arXiv.2302.09319>

⁴⁴ Ng et al i. m.

⁴⁵ Carolus et al. i. m.

Az elmúlt években a méréseknek köszönhetően a tudományos vizsgálat már az MI-jártassághoz kapcsolódó tipológiák felállítását is elkezdte. A felsőoktatás vonatkozásában Bewersdorff és munkatársai 2024-es kutatását érdemes említenünk. Ők úgy vizsgálták a MI-jártasságot egyetemi hallgatók (1465 diák az USA-ból, Nagy-Britanniából és Németországból) körében, hogy azt az énhatékonysággal kapcsolták össze.⁴⁶ Felmérték a hallgatók mesterséges intelligencia iránti érdeklődését, attitűdjeiket és a MI-használatukat. Ezek alapján három csoportot azonosítottak: az MI támogatói (*AI Advocates*), az óvatos kritikusok (*Cautious Critics*) és pragmatikus megfigyelők (*Pragmatic Observers*) csoportját. Szerintük olyan MI-oktatási stratégiára van szükség, amely figyelembe veszi a különböző hallgatói típusok eltérő igényeit. A jövő évek vizsgálatai, a MI-jártasság mérései valószínűleg ebbe az irányba fognak folytatódni.

6. Összegzés

A tanulmány arra vállalkozott, hogy a mesterséges intelligenciához kapcsolódó kompetenciaköröket bemutassa. Ennek elméleti kontextusát úgy teremtette meg, hogy először áttekintette a 2000-es évektől megjelent, a hálózati, digitális kommunikációhoz kapcsolódó kompetencialeírásokat, különös tekintettel Wilson és munkatársai által 2011-ben az UNESCO számára kidolgozott munkájára. Innen lépett tovább a bemutatás a mesterséges intelligencia sikeres és eredményes használatához szükséges kompetenciák bemutatására. A tanulmány ezen szövegrészében is az UNESCO által kidolgozott keret kapta a legnagyobb figyelmet, mind a tanulók, mind a tanárok irányába megfogalmazott kompetenciák tekintetében. Ezek a keretrendszer a „*AI competency framework for students – AILFS*” és az „*AI competency framework for teachers – AILFT*” nevet viselik. Ugyanakkor ezen modell mintegy kritikájaként Annapureddy és munkatársai felhasználó-centrikus, empirikus kutatásokon alapuló 12 elemű keretrendszerét is bemutatta a tanulmány. Az írás befejező részében pedig két, a MI-műveltség mérésére alkalmazott eszközt mutattam be, az egyik Ng és munkatársai által kidolgozott *AILQ – AI Literacy Questionnaire*, a másik pedig Carolus és munkatársaihoz köthető *MAILS* volt. A dolgozat limitációi közé tartozik, hogy nem tudta széleskörűen bemutatni az MI-jártassággal foglalkozó kutatásokat, ezek rendkívüli sokasága miatt, csak a tudományosan legelfogadottabbakra tért ki.

⁴⁶ Arne Bewersdorff – Marie Hornberger – Claudia Nerdel – Daniel S. Schiff: AI advocates and cautious critics: How AI attitudes, AI interest, use of AI, and AI literacy build university students' AI self-efficacy. *Computers and Education: Artificial Intelligence*, Vol. 8. (2025).

A MESTERSÉGES INTELLIGENCIA SZÓLÁSSZABADSÁGA ÉS TARTALOMMODERÁLÁSI GYAKORLATA A KÖZÖSSÉGI MÉDIÁBAN A DSA ÉS AZ MI RENDELET VONATKOZÁSÁBAN

Gosztonyi Gergely* – Gyetván Dorina** – Kovács Andrea***

1. Bevezetés

Bár Gartner technológiák felfutását vizsgáló úgynevezett Hype Cycle görbéjén a generatív mesterséges intelligencia már 2024 augusztusában a ‘csalódási fázisba’ került,¹ egyre több szoftverben tapasztaljuk valamilyen mesterséges intelligencia megjelenését. Ez a trend nem kerülheti el a közösségi média platformokat sem, a Facebook és az Instagram mögött álló Meta például saját mesterséges intelligencia modellt fejleszt Llama néven, amelyet a felhasználói bátran testre szabhatnak.² Emellett a Meta nyíltan vállaltan használ mesterséges intelligenciát a tartalmak szűrésére,³ ahogyan az X / Twitter, valamint a TikTok is hasonló módszert alkalmaz.

A mesterséges intelligencia – a közvélemény számára hirtelennek tűnő – felkapottsága a jogalkotókat sem hagyta hidegen. A probléma ugyanakkor az, hogy – Pázmándi Kinga szavaival – „a tolmérce egyik végén a jog által biztosítani kívánt védelmi célok [...] állnak, a másik végén a társadalmilag jól hasznosuló innováció, a

* Egyetemi tanár, Eötvös Loránd Tudományegyetem, Állam- és Jogtudományi Kar, gosztonyi@ajk.elte.hu, ORCID: <https://orcid.org/0000-0002-6551-1536>

** PhD hallgató, Eötvös Loránd Tudományegyetem Állam- és Jogtudományi Doktori Iskola, dorina.gyetvan@gmail.com, ORCID: <https://orcid.org/0000-0001-5361-0011>

*** PhD hallgató, Eötvös Loránd Tudományegyetem Állam- és Jogtudományi Doktori Iskola, ankovacs@student.elte.hu, ORCID: <https://orcid.org/0000-0003-4970-3551>

¹ 2024 Hype Cycle for Emerging Technologies Highlights Developer Productivity, Total Experience, AI and Security. *Gartner*, 2024. szeptember 21. <https://tinyurl.com/5n8jx2uz>

² <https://www.llama.com>

³ Meta: Meta’s New AI System to Help Tackle Harmful Content. *Meta*, 2021. december 8. <https://tinyurl.com/4w9xup4r>



»tolómérce értékeinek beállítását« pedig állandó [...] bizonytalanság kíséri.⁴ A tanulmány megírásának időpontjában az elmúlt egy éves időtartamban az Európai Unióban (a továbbiakban: EU) elfogadták a mesterséges intelligencia rendeletet⁵ (a továbbiakban: MI rendelet) és hozzá kapcsolódóan a felelőségi irányelv tervezetét,⁶ az Egyesült Államok korábbi elnöke, Joseph Biden pedig kibocsátott egy utasítást a biztonságos és megbízható mesterséges intelligencia fejlesztéséről és használatáról (a továbbiakban: E.O. 14110).⁷ Az utasítást Donald J. Trump a mesterséges intelligencia fejlesztés akadályának tekintette és több más korábbi elnöki rendelettel együtt 2025 elején visszavonta,⁸ egyúttal utasítva az érintett szervezeteket, hogy azonnali hatállyal vizsgálják felül korábbi lépéseiket.⁹

Bár általánosan elfogadott definíció nem létezik a szakirodalomban,¹⁰ mind az EU, mind pedig az Egyesült Államok meghatározza, hogy mit ért mesterséges intelligencia alatt. Az EU rendelete szerint a mesterséges intelligencia olyan szoftver, amely mögött meghatározott technológia áll (gépi tanulás) vagy meghatározott megközelítéssel került kifejlesztésre (logikai és tudás alapú, illetve statisztikai), és amely képes a környezetét befolyásolni.¹¹ Az E.O. 14110 és az azt váltó elnöki rendelet, a US Code definíciójára hivatkozik, mely szerint a mesterséges intelligencia egy gépi rendszer, amely előrejelzéseket ad, ajánlásokat tesz vagy döntéseket hoz, amelyhez a bemeneti adatokat

⁴ Pázmándi Kinga: Digitalizáció, technológiai fejlődés, jogi paradigmák. *Gazdaság és Jog*, 2018/12. 11. <https://doi.org/10.21637/GT.2018.01.02>

⁵ Az Európai Parlament és a Tanács (EU) 2024/1689 rendelete (2024. június 13.) a mesterséges intelligenciára vonatkozó harmonizált szabályok megállapításáról, valamint a 300/2008/EK, a 167/2013/EU, a 168/2013/EU, az (EU) 2018/858, az (EU) 2018/1139 és az (EU) 2019/2144 rendelet, továbbá a 2014/90/EU, az (EU) 2016/797 és az (EU) 2020/1828 irányelv módosításáról (a mesterséges intelligenciáról szóló rendelet), PE/24/2024/REV/1, HL L, 2024/1689, 2024.7.12. Varju Márton: A mesterséges intelligencia szabályozása az Európai Unióban: szakpolitika, szabályozási stratégia és jog. In: Mezei Kitti (szerk.) *Kockázatok és lehetőségek a mesterséges intelligencia jogi szabályozásában*. Budapest, Wolters Kluwer, 2025. 33–50.

⁶ Javaslat az Európai Parlament és a Tanács irányelve a szerződésen kívüli polgári jogi felelősségre vonatkozó szabályoknak a mesterséges intelligenciához való hozzáigazításáról (a mesterséges intelligenciával kapcsolatos felelősségről szóló irányelv) COM(2022) 496 final 2022/0303 (COD). Ki kell ugyanakkor emelni, hogy 2025 tavaszán az Európai Bizottság *ad acta* kívánja helyezni a felelőségi irányelv tervezetét. Isabel Marques da Silva: Vajon az MI-specifikus felelőségi szabályok visszavonása jogorvoslat nélkül hagyja a károsultakat? *Euronews*, 2025. március 4. <https://tinyurl.com/2xnkt38f>

⁷ Executive Order (E.O.) 14110 on Safe, Secure and Trustworthy Development and Use of Artificial Intelligence, 88 FR 75191.

⁸ The White House: Initial Rescissions of Harmful Executive Orders and Actions. *whitehouse.gov*, 2025. január 20. <https://tinyurl.com/4wuhbe8h>

⁹ The White House: Removing Barriers to American Leadership in Artificial Intelligence. *whitehouse.gov*, 2025. január 23. <https://tinyurl.com/2u6zf73x> Sec. 5.

¹⁰ Necz Dániel: A mesterséges intelligencia felhasználásával történő adatkezelések egyes sajátos szempontjai. *Acta Humana*, 2023/3. 98. <https://doi.org/10.32566/ah.2022.3.4>; Ambrus István: A mesterséges intelligencia és a büntetőjog. *Állam- és Jogtudomány*, 2020/4. 6–9.

¹¹ MI rendelet 3. cikk 1.

modellekké absztrahálja.¹² Megjegyzendő, hogy a definíció eredeti kontextusát tekintve a kereskedelmi szabályok között helyezkedik el a National AI Initiative cím alatt.

A jogalkotókkal szemben a szakirodalom már szentimentálisabb, sokkal inkább az emberi gondolkodáshoz való hasonlóságra vagy az attól való különbségekre koncentrálna.¹³ Amikor a szakirodalom a mesterséges intelligencia szólásszabadságát – és ehhez kapcsolódóan a személyiség kérdését – vizsgálja, akkor általában a kevésbé technikai, sokkal inkább emberi vonások kerülnek előtérbe, amellyel jelen tanulmány első része foglalkozik. A tanulmány a továbbiakban vizsgálja a tartalommoderálás általános kérdéseit és szabályozását az EU-ban, illetve a mesterséges intelligencia – pozitív és negatív – szerepét és esélyeit a tartalommoderálási gyakorlatban.

2. A mesterséges intelligencia szólásszabadsága

Egyes mesterséges intelligenciák esetében láthattunk már példát arra, hogy alkotójától elkülönülten valamilyen joggal ruházzák fel, mint Sophia¹⁴ és Mirai esetében.¹⁵ A mesterséges intelligencia egyre elterjedtebb alkalmazása miatt azonban leggyakrabban a felelősségi kérdésekkel foglalkozik a szakirodalom,¹⁶ ugyanakkor – amennyiben a fizikai dimenziótól eltekintünk – a robotjog már régóta foglalkozik az ember által megteremtett, mesterségesen létrehozott, intelligens gépek lehetséges jogaival, amelyekbe beletartozik a kommunikációhoz való jog.¹⁷

A mesterséges intelligencia és a szólásszabadság kapcsán érdemes megvizsgálni, hogy ki a szólásszabadság alanya. Potenciális alany lehet:

- (i) a mesterséges intelligencia, vagy
- (ii) egy mögöttes alany, közülük pedig

¹² 15 USC 9401: Definitions (Pub. L. 116–283, div. E, §5002, Jan. 1, 2021, 134 Stat. 4523.); E.O. 14110 Section 3. (b). Theodore S. Boone: An examination of certain key features of the new White House Executive Order on Artificial Intelligence. *Corvinus Law Papers*, 2024/1. 1–19.

¹³ Az erős és gyenge mesterséges intelligenciák jellemzőiről ld. összefoglalóan: Kovács Andrea: A mesterséges intelligencia jogi megítélésének egyes kérdései. *Themis*, 2024/1. 37–40. <https://doi.org/10.55052/themis.2024.1.36>

¹⁴ Sophia például állampolgárságot kapott. Molnár Balázs: Személy-e a mesterséges intelligencia? In: Varga János – Csiszárik-Kocsir Ágnes – Garai-Fodor Mónika (szerk.): *Vállalkozásfejlesztés a XXI. században 2023. II. kötet. A jelen kor gazdasági kihívásainak és társadalmi változásainak interdiszciplináris megközelítései*. Budapest, Óbudai Egyetem Keleti Károly Gazdasági Kar, 2023. 280–283.

¹⁵ Mirai az üzemeltető vállalatától függetlenül letelepedési engedélyt kapott. Nagy Teodóra: A jövő kihívásai: robotok és mesterséges intelligencia az alapjogi jogalanyiság tükrében. *MTA Law Working Papers*, 2020/6. 13.

¹⁶ Ld. pl. Stefán Ibolya: A mesterségesintelligencia-rendszerek felelősségi kérdései az uniós dokumentumok és a magyar szabályozás tükrében. *Publicationes Universitatis Miskolcensis, Sectio Juridica et Politica*, 2022/2. 364–387. <https://doi.org/10.32978/sjp.2022.030>; Csítei Béla: Az önvezető gépjárművek és a polgári jogi kárfelelősség. In: Glavanits Judit (szerk.): *A gazdasági jogalkotás aktuális kérdései*. Budapest, Dialóg Campus, 2019. 25–26.

¹⁷ Yurii Vadymovych Sheliashenko: Artificial Personal Autonomy and Concept of Robot Rights. *cyberleninka.ru*, 2017. április 12., 17–21. <http://dx.doi.org/10.20534/EJLPS-17-1-17-21> ; Klein Tamás: Robotjog. In: Klein Tamás – Tóth András (szerk.): *Technológiajog – Robotjog – Cyberjog*. Budapest, Wolters Kluwer Kft., 2018. 179–215. <https://doi.org/10.55413/9789632958293>

- a) a programozó,
- b) az üzemben tartó / üzemeltető, vagy
- c) a használó.

A lehetséges alanyok közül maga a mesterséges intelligencia alanyiséga a legvitatottabb. A szólásszabadság alanyaként való elismerésének lehetősége nem idegen a szakirodalomtól,¹⁸ ehhez különböző szempontokat határoznak meg az egyes szerzők, amelyek az egészen technikai, viszonylag egyszerűen meghatározható feltételektől az absztrakt feltételekig terjednek. Egyszerűen meghatározható szempont például, hogy a szólásszabadság elismeréséhez elegendő az is, hogy a kimeneti eredmény indeterminált, nem megjósolható.¹⁹ Hasonló álláspontot képvisel Tim Wu is, aki szerint akkor beszélhetünk személyiségről, ha (1) képes koncepcionális gondolkodásra és (2) képes kifejezni a kialakult véleményét.²⁰ Haladva az absztraktabb feltételek felé, Sheliazhenko szerint a szólásszabadság elismeréséhez azt szükséges vizsgálni, hogy rendelkezik-e a vizsgálat tárgya olyan társadalmilag komplex funkcionalitással, amely az emberi autonómia határait súrolja,²¹ Schwitzgebel és Garza szerint pedig a megfelelő szociális státusz is elegendő lehet.²² Érdeemes azonban megemlíteni, hogy a mesterséges intelligencia is felruházható jogi személyiséggel – így bizonyos jogokkal – is, a jogi személyiség megadása pedig a jogalkotó diszkrecionális döntése és mint ilyen, akár praktikussági szempontokat is követhet.²³

Más szemszögből nézve azonban a mesterséges intelligencia nem önmagában való, mindig áll mögötte valamilyen emberi tevékenység,²⁴ így amennyiben nem a mesterséges intelligenciának tulajdonítjuk a szólásszabadságot, úgy a mögöttes személyeket szükséges vizsgálnunk.²⁵ A programozó, az üzemben tartó/üzemeltető és a használó

¹⁸ Bert-Jaap Koops – Mireille Hildebrandt – David-Olivier Jaquet-Chiffelle: Bridging the Accountability Gap: Rights for New Entities in the Information Society? *SSRN*, 2010. július 23. 555–559. *Minnesota Journal of Law, Science & Technology*, Vol. 11., No. 2. (2010) 497–561.

¹⁹ Marek Świerczyński – Zbigniew Więckowski: Statut Jednolity Sztucznej Inteligencji. *Zeszyty Prawnicze*, Vol. 23., No. 1. (2023) 220. <https://doi.org/10.21697/zp.2023.23.1.09>

²⁰ Tim Wu: Machine Speech. *University of Pennsylvania Law Review*, Vol. 161., No. 6. (2013) 1495., 1503. Ezzel kapcsolatban ld. még: *Autronic AG v Switzerland*, no. 12726/87 1990. május 22-i ítélet, 47.

²¹ Sheliazhenko i. m. 19.

²² Eric Schwitzgebel – Mara Garza: A Defense of the Rights of Artificial Intelligences. *Midwest Studies In Philosophy*, Vol. 39., No. 1. (2015) 102–103. <https://doi.org/10.1111/misp.12032>

²³ Beatriz A. Ribeiro et al.: Metacognition, Accountability and Legal Personhood of AI. In: Henrique Sousa Antunes et al. (szerk.): *Multidisciplinary Perspectives on Artificial Intelligence and the Law*. Cham, Springer, 2024. 174., 181–182. https://doi.org/10.1007/978-3-031-41264-6_9; Visa Aj Kurki: *A Theory of Legal Personhood*. Oxford, Oxford University Press, 2019. 177., 179. <https://doi.org/10.1093/oso/9780198844037.001.0001>; Keserű Barna Arnold: A mesterséges intelligencia néhány magánjogi aspektusáról. In: Glavanits Judit (szerk.): *A gazdasági jogalkotás aktuális kérdései*. Budapest, Dialóg Campus, 2019. 116. A témáról bővebben ld. Kovács i. m.

²⁴ Margot E. Kaminski – Meg Leta Jones: Constructing AI Speech. *The Yale Law Journal Forum*, Vol. 133., No. 1. (2024) 1226.

²⁵ Federico Gustavo Pizzetti: Embryos, Organoids and Robots: ‘legal subjects’? *BioLaw Journal*, Vol. 9., No. 1. (2021) 348–349. <https://doi.org/10.15168/2284-4503-755>; Matthew Hines: I Smell a Bot:

esetében ezek a szerepkörök átjárhatók, közöttük nem húzódik éles határ. Moderálás tekintetében példákkal szemléltetve a lehetséges eseteket:

- a) a programozó, az üzemben tartó/üzemeltető és a használó megegyezik. Ebben az esetben ugyanaz a vállalat programozza, üzemelteti (elhárítja a hibákat, 1-2-n szintű támogatást nyújt) és használja (promptolja, paraméterezi, döntéseket hoz a kimenet alapján);
- b) a vállalat a mesterséges intelligencia kapcsán egy másik vállalatot bíz meg a fejlesztéssel, de azt már ő maga üzemelteti és használja;
- c) a cég csak megvásárolja és használja a mesterséges intelligenciát, míg a programozást és az üzemeltetést más végzi.

Míg az a) esetben a helyzet egyértelmű, a jogok és a felelősség is a vizsgált vállalat-hoz kapcsolódik, addig a c) esetben már az is kérdés, hogy mennyiben egy kompakt, úgynevezett dobozos szoftverről, és mennyiben egyedi fejlesztésű megoldásról van szó. Az üzemeltető szerepkörét tekintve iránymutatásként használható az MI rendelet, amely e szerepkörebe tartozónak tekinti a szolgáltatót, az alkalmazót, a meghatalmazott képviselőt, az importőrt és a forgalmazót is, amennyiben a mesterséges intelligencia veszélyes üzem.²⁶

Ezen szereplők közül azonban feltehetően azt fogjuk beszélőnek tekinteni, aki a tartalom generálásához (vagy jelen esetben moderálásához) szükséges paramétereket meghatározta.²⁷ Amiben azonban a programozó, az üzemeltető és használó szerepköre megegyezik, hogy természetes és/vagy jogi személyiséggel rendelkeznek, amelyek rendelkeznek szólásszabadsággal. A közösségi média szolgáltatók tekintetében a saját fejlesztés miatt a mesterséges intelligencia tevékenysége feltehetően a szolgáltatóknak lesz betudható, különösen arra tekintettel, hogy azokat a szabályzatokat és irányelveket is a szolgáltató írja, amelyeket végre kell hajtani.²⁸ Ezt az álláspontot látszik követni a 3rd Circuit ítélete az Anderson v. TikTok ügyben.²⁹ Szemben az alsóbb fokú bíróság ítéletével bíró értelmezésében – a Moody v. NetChoice³⁰ Legfelső Bírósági ítéletre alapozva – a TikTok felelősségre vonható az algoritmusait által való ajánlásért és összeállításért, mint saját kifejező produktumáért.³¹ Kiemelendő azonban, hogy az ügyben végig algoritmusról van szó, nem pedig mesterséges intelligenciáról. A Moody v. NetChoice-ítélet ugyan felveti a mesterséges intelligencia alkalmazásának kérdését,

California's S.B. 1001, Free Speech, and the Future of Bot Regulation. *Houston Law Review*, Vol. 57., No. 2. (2019) 408.

²⁶ MI rendelet 3. cikk, 3–7.

²⁷ Hines i. m. 420.

²⁸ Kylie Robison: Inside the Shifting Plan at Elon Musk's X to Build a New Team and Police a Platform 'so Toxic It's Almost Unrecognizable'. *Fortune*, 2024. február 7. <https://tinyurl.com/37x2hu3c>

²⁹ Anderson v. TikTok Inc, No. 22-3061 (3d Cir. 2024).

³⁰ Moody v. NetChoice, LLC. 603 U.S. 707 (2024).

³¹ Anderson v. TikTok, 184–186.

de mindössze annyi megjegyzést tesz, hogy ilyen eszközök alkalmazása alkotmányos jelentőségű lehet.³²

Más szerzők azonban nem foglalkoznak a jogalanyiség kérdésével, más igazolást választanak a mesterséges intelligencia kimenetének védelméhez, részben azért, mert a mesterséges intelligencia esetében – épp annak kiszámíthatatlansága miatt – előfordulhat, hogy tevékenysége nem embernek betudható, nem vezethető vissza könnyedén emberi tevékenységre.³³ Ezen érvek többek között:

1. Nincs olyan doktrína, amely kifejezetten kizárná a mesterséges intelligenciát az Első Alkotmánykiegészítés hatálya alól.³⁴
2. A mesterséges intelligencia kimenete tartalmazhat olyan elemeket, amelyek az emberek számára érdekesek, fontosak, vagy tartalmazhat olyan közléseket, amelyeket egyébként a szólásszabadság véd.³⁵ Ezáltal pedig ezek részt vesznek a demokratikus vitában és az információk terjesztésében.³⁶ Ezen érv részeként kiemelendő a hallgatóság joga az információhoz jutáshoz, amely a szólásszabadság része.³⁷

A mesterséges intelligencia által generált üzenetek esetében következményekkel jár az is, ha elfogadjuk, hogy a szólásszabadság hatálya alá tartozik, és az is, ha nem. Ha nem tartozik a hatálya alá pusztán azért, mert nem közvetlenül embertől származik, akkor azzal egyébként védett beszéd cenzúrázható.³⁸ Ha viszont nem zárható ki a szólásszabadság hatálya alól, akkor problémát okozhat az egyébként a beszélő szándékától függően nem védett beszéd megítélése és kérdésessé válik a felelősségre vonás.³⁹ A közösségi média esetében – amennyiben ténylegesen saját fejlesztésű mesterséges intelligenciát használnak tartalommoderálásra – a helyzet egyértelmű: a fejlesztés és a paraméterezés (végrehajtandó szabályzatok megalkotása és a végrehajtás módja) is egy kézben van.⁴⁰

³² Moody v. NetChoice.

³³ Lynne Higby: Navigating the Speech Rights of Autonomous Robots in a Sea of Legal Uncertainty. *Journal of Technology Law & Policy*, Vol. 26., No. 1. (2021) 33.

³⁴ Toni Massaro – Helen Norton: Siri-Ously? Free Speech Rights and Artificial Intelligence. *Northwestern University Law Review*, Vol. 110., No. 5. (2016) 1185.

³⁵ Koops–Hildebrandt–Jaquet–Chiffelle i. m. 555–559.; Massaro–Norton i. m. 1192.; James Garvey: Let’s Get Real: Weak Artificial Intelligence Has Free Speech Rights. *Fordham Law Review*, Vol. 91., No. 3. (2022) 953., 991.

³⁶ Kaminski–Jones i. m. 1231–1242. A szerzők azt is kifejtik, hogy vannak olyan esetek, amikor fontos szerepet tölt be a beszélő személye, illetve szándéka (például rágalmazás esetén). A magyar Alkotmánybíróság gyakorlatában a nyilvános társadalmi kommunikációban való részvételben látja a szólásszabadság hatályát. Török Bernát: *Szabadon szólni, demokráciában. A szólásszabadság magyar doktrínája az amerikai jogirodalom tükrében*. Budapest, HVG-ORAC, 2018.

³⁷ Emberi Jogok Európai Egyezménye 10. cikk.

³⁸ Higby i. m. 47.; Wu i. m. 1521.

³⁹ Higby i. m. 48.

⁴⁰ Robison i. m.

3. A tartalommoderálás általános kérdései és szabályozása az EU-ban

3.1. A tartalommoderálás és a DSA, különös tekintettel az átláthatósági követelményekre

2021 márciusában Mark Zuckerberg úgy nyilatkozott kongresszusi meghallgatása során, hogy a gyűlöletbeszédet tartalmazó tartalmak 95%-át, míg – legjobb tudomása szerint – a terrorista tartalmat megtestesítő tartalmak 98-99%-át már mesterséges intelligencia azonosítja, és nem emberek.⁴¹ A Facebook, az X/Twitter és a többi közösségi média platform moderálási gyakorlatának, illetve elveinek szabályozása világszerte nagy kihívást jelent. A digitális szolgáltatások egységes piacáról szóló 2022/2065 rendelet⁴² (a továbbiakban: DSA) elsősorban a közvetítő szolgáltatókra, így az online platformokra vonatkozó szabályozást tartalmaz, azonban a tartalommoderálásra vonatkozó rendelkezések, illetve az algoritmikus átláthatósági és elszámoltathatósági követelmények folytán kiegészíti a többi uniós, mesterséges intelligencia szabályozására irányuló törekvést.⁴³

A DSA e körben választ kísérel adni a tartalommoderálás egyes problémás kérdéseire is, célkitűzései között szerepel az átláthatóság és elszámoltathatóság biztosítása,⁴⁴ illetve az online platformot üzemeltető szolgáltatók tartalommoderálási döntéseinek hatékony felülvizsgálhatósága.⁴⁵

Ugyan a DSA már meghatározza a tartalommoderálás fogalmát, a tartalommoderálásnak korábban nem volt – és a szakirodalmat tekintve továbbra sincs – tudományos konszenzuson alapuló meghatározása, az inkább egyfajta gyűjtőfogalomként fogható fel.⁴⁶ Összefoglalóan – tág értelemben – úgy határozhatjuk meg, mint amely magában foglalja a közösségi média szolgáltatók minden alapjogi kérdéseket felvető olyan eljárását, amely minden esetben felhasználói tartalmat érint az ahhoz való hozzáférhetőség emberi erőforrás vagy mesterséges intelligencia általi korlátozásával, amelynek azonban nem előfeltétele az érintett tartalom jogellenessége – mivel elsődleges „jogalapját” a jogszabályok helyett a platformok szerződéses feltételei között kell keresni –, és ame-

⁴¹ House Energy and Commerce Subcommittee on Communications and Technology Disinformation Nation: *Social Media's Role in Promoting Extremism and Misinformation*. 117th Congress. Testimony of Mark Zuckerberg, 2021. március 25. <https://tinyurl.com/52by7xzn>

⁴² Az Európai Parlament és a Tanács (EU) 2022/2065 rendelete (2022. október 19.) a digitális szolgáltatások egységes piacáról és a 2000/31/EK irányelv módosításáról (digitális szolgáltatásokról szóló rendelet), PE/30/2022/REV/1, HL L 277., 2022.10.27., 1–102.

⁴³ Mayer Brown: EU Digital Services Act's Effects on Algorithmic Transparency and Accountability. *Lexology*, 2023. március 27., <https://tinyurl.com/8wsa82x2>

⁴⁴ DSA 49. preambulumbekzdés.

⁴⁵ DSA 44. és 109. preambulumbekzdés.

⁴⁶ Gosztanyi Gergely: *Censorship from Plato to Social Media. The Complexity of Social Media's Content Regulation and Moderation Practices*. Cham, Springer, 2023. 7–19. https://doi.org/10.1007/978-3-031-46529-1_2

lyet a platform általában a saját gazdasági vagy társadalmi agendája⁴⁷ vagy a közvetítő szolgáltatói felelősség alóli mentesülés érdekében alkalmaz, így hiányzik a mérlegelés és a döntéshozatal átláthatósága. A motivációk közül a saját gazdasági érdekből történő moderálás (azért, hogy a felhasználók ne találkozzanak sértő, felháborító vagy zavaró tartalmakkal) annál meghatározóbb, minél gyakoribbak a felhasználók közötti véleménycserék, és minél élénkebb a kommunikáció.⁴⁸

Ezzel szemben a DSA szerinti tartalommoderálás a közvetítő szolgáltató olyan automatizált vagy nem automatizált tevékenysége, amely különösen a szolgáltatás igénybe vevője által közzétett jogellenes tartalom vagy a közvetítő szolgáltató szerződési feltételeivel összeegyeztethetetlen információ észlelésére, azonosítására és kezelésére szolgál, ideértve az ilyen jogellenes tartalom vagy információ elérhetőségét, láthatóságát és hozzáférhetőségét érintő intézkedéseket, például annak hátra sorolását, demonetizálását, az ahhoz való hozzáférés megszüntetését vagy annak eltávolítását, vagy a szolgáltatás igénybe vevője általi információközlés lehetőségét érintő intézkedéseket, például a fiókja megszüntetését vagy felfüggesztését.⁴⁹ A fenti két meghatározást összevetve láthatjuk, hogy nincs teljes tartalmi azonosság a két meghatározás között. A DSA nem teszi fogalmi elemmé az alapjogi vonatkozást, holott ezen döntések mindig – legalább – a felhasználók szólásszabadságának korlátozásával járnak. A másik legfontosabb különbség a meghatározások között, hogy míg a DSA alapján a tartalommoderálás a moderálással érintett tartalom felderítésére vonatkozó eljárást is a moderálás részének tekinti, addig a felderítés aligha tekinthető moderálásnak abban az értelemben, hogy a tartalommal kapcsolatos tényleges beavatkozás hiányában nem érint alapvető jogokat. A jelen tanulmány céljából, annak alapjogi szemléletére tekintettel a szűkebb, DSA-tól eltérő, beavatkozás nélküli felderítés aktusát nem tekintjük moderálásnak.

A platformok nagyrészt automatizált döntéshozatalán alapuló moderálási döntéseinek ellensúlyozására, illetve az alapvető jogok korlátozásának garanciákkal történő védelme⁵⁰ érdekében a DSA alapvetően öt eszközt tartalmaz: (i) tartalommoderálási tájékoztató – eljárási szabályzat; (ii) átláthatósági jelentések; (iii) indokolási kötelezettség; (iv) belső panaszkezelési rendszer és személyzeti felügyelet és (v) peren kívüli vitarendezés.

Ad (i) A 14. cikk kötelezi a közvetítő szolgáltatókat arra, hogy szerződéses feltételeik könnyen hozzáférhető és géppel olvasható formátumban olyan világos, egyszerű, érthető és felhasználóbarát tájékoztatást tartalmazzanak a tartalommoderálásra vonatkozóan, amely magában foglalja valamennyi alkalmazott szabályra, eljárásra, intézkedésre és eszközre – beleértve az algoritmikus döntéshozatalt és az emberi felülvizsgá-

⁴⁷ Koltay András: A social media platformok jogi státusa a szólásszabadság nézőpontjából. *In Medias Res*, 2019/1. 42.

⁴⁸ Uo. 47.

⁴⁹ DSA 3. cikk t).

⁵⁰ DSA 9. preambulumbekzdés.

latot –, valamint a belső panaszkezelési rendszerük eljárási szabályzatára vonatkozó információkat.⁵¹

Ad (ii) A mesterséges intelligencia alaptulajdonságai között említendő az összetettség, az adatoktól való függés,⁵² az eredendő átláthatatlanság,⁵³ a kiszámíthatatlanság és a kiismerhetetlenség,⁵⁴ amelyekből logikusan adódik, hogy ezek ellensúlyozására⁵⁵ az MI szabályozásának általános kialakítása kapcsán az egyik fő erkölcsi eredetű, illetve alapjogi célkitűzés a magyarázhatóság, vagyis a működés átláthatósága.⁵⁶ A mesterséges intelligencia átláthatatlansága továbbá alkalmas arra, hogy például közösségi média felhasználók alapvető eljárásjogi jogainak érvényesülését akadályozza, így különösen a hatékony jogorvoslathoz, a tisztességes eljáráshoz való jog érvényesülését.⁵⁷

Ugyanakkor az MI további alapjellemzője az autonóm magatartás is,⁵⁸ amely miatt például az átláthatóság biztosítása komoly kihívást jelent különösen gépi tanulásra tekintettel.⁵⁹ Ezen problémát erősíti az ún. „feketedoboz jelenségnek” (*black box effect*) nevezett előfeltevés is, amely alapvetően az MI rendszerek azon jellemzőjére utal, hogy az autonóm mesterséges intelligencia rendszerek emberek (felhasználók) által megismerhetetlenül, értelmezhetetlenül működnek.⁶⁰ Ezen gondolatmeneten továbbhaladva azzal a paradox helyzettel találkozhatjuk szembe magunkat (információ-túlterhelés vagy átláthatósági paradoxon⁶¹), hogy teljes transzparencia, azaz a pontos programozási kód vagy algoritmus nyilvánosságra hozatala sem képes a transzparenciával elérni kívánt cél megvalósítására, hiszen a felhasználók megfelelő ismereteik hiányában nem képesek, vagy pedig csak aránytalan idő- és energiaráfordítással képesek megérteni a rendszerek működését, amely így pontosan az ellenkező hatást éri el: inkább félrevezeti a felhasználókat, mint felvilágosítja őket a rendszer megbízhatóságáról. Továbbá a

⁵¹ Martin Husovec: Will the DSA Work? In: Joris van Hoboken et al. (szerk.): *Putting the DSA into Practice. Enforcement, Access to Justice and Global Implications*. Berlin, Verfassungsblog gGmbH, 2023. 26. <https://doi.org/10.17176/20230208-093135-0>

⁵² European Commission, Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts (a továbbiakban: EC Proposal) 3.5.

⁵³ Papadouli Vasiliki: Transparency in Artificial Intelligence: A Legal Perspective. *Journal of Ethics and Legal Technologies*, Vol. 4., No. 1. (2022) 25.

⁵⁴ Tóth András: A mesterséges intelligencia szabályozásának paradoxonja és egyes jogi vonatkozásainak alapvető kérdései. *Infokommunikáció és Jog*, 2019/2. 3.

⁵⁵ Vasiliki i. m. 25.

⁵⁶ Mezei Kitti: A mesterséges intelligencia jogi szabályozásának aktuális kérdései az Európai Unióban. *In Médias Res*, 2023/1. 55. <https://doi.org/10.59851/imr.12.1.4>

⁵⁷ Uo. 62.

⁵⁸ EC Proposal 3.5.

⁵⁹ Kovács Zoltán Balázs: Nincs új a nap alatt, avagy figyelemmel az eddigi Európai Unió hatósági gyakorlatra is, választ ad a GDPR a mesterséges intelligencia használata kapcsán felvetődő adatvédelmi kérdésekre? *Szecska Ügyvédi Iroda*, 2024. június 18., 8. <https://tinyurl.com/3vtuzs3j>

⁶⁰ Alexander J. Wulf–Ognyan Seizov: Artificial Intelligence and Transparency: A Blueprint for Improving the Regulation of AI Applications in the EU. *European Business Law Review*, Vol. 31., No. 4. (2020) 619. <https://doi.org/10.54648/EULR2020024>

⁶¹ Vasiliki i. m. 32.

transzparencia hétköznapi értelemben azt is jelenthetné, hogy a mesterséges intelligenciára vonatkozó tudásanyag szisztematikusan átadásra kerülne egyik érdekelt féltől a másiknak,⁶² amely azonban azt eredményezheti, hogy az MI rendszerekre vonatkozó nem megfelelő mértékű és módon történő átláthatósági követelmények előírása aránytalan sérelmet jelentene a nem nyilvános know-how és üzleti információk (üzleti titkok) felfedésére kötelezéssel.⁶³

A DSA az automatizált döntéshozatalán alapuló moderálási döntéseinek ellensúlyozására átláthatósági jelentési kötelezettségeket ír elő a szolgáltatókra vonatkozóan.⁶⁴ Ezen kötelezettség három szinten kerül meghatározásra kumulatív jelleggel,⁶⁵ egyrészt valamennyi közvetítő szolgáltatóra vonatkozó minimális tartalmi elemek keretében,⁶⁶ másrészt az online platformot üzemeltető szolgáltatókra vonatkozóan,⁶⁷ és harmadrészt pedig az online óriásplatformot vagy nagyon népszerű online keresőprogramot üzemeltető szolgáltatókra irányadóan.⁶⁸

Az óriásplatformnak minősülő, legnépszerűbb közösségi szolgáltatók az átláthatósági jelentésekben az automatizált döntéshozatallal kapcsolatosan az alábbi információkat kötelesek feltüntetni:

- a) érdemi és érthető tájékoztatást a szolgáltató saját kezdeményezésére végzett tartalommoderálásról, beleértve az automatizált eszközök használatát is;⁶⁹
- b) arra vonatkozó információt, hogy automatizált eszközöket alkalmaz tartalommoderálás céljából, beleértve a minőségi leírást, a pontos célok meghatározását, az e célok eléréséhez használt automatizált eszközök pontosságára és lehetséges hibaarányára vonatkozó mutatókat, valamint az alkalmazott biztosítékokat;⁷⁰
- c) a (b) pontban írt automatizált eszközökre vonatkozóan pontossági mutatókra és kapcsolódó információkra kiterjedő tájékoztatást.⁷¹

Ad (iii) A DSA megalkotásának egyik fő célkitűzése a hatékony jogorvoslati mechanizmusok hozzáférhetőségének javítása volt.⁷² Azért, hogy a felhasználók hatékonyan tudjanak élni jogorvoslati jogukkal, valamennyi tárhelyszolgáltató köteles a döntését (nem csak a tartalmak törlése, hanem például a demonetizáció vagy a tartalom hátra sorolása esetén is) megfelelően megindokolni, ha moderálást végez, ugyanis az indokolásnak alapvető jelentősége van a jogorvoslati jog gyakorlása során. A közösségi média-

⁶² Reid Blackman – Beena Ammanath: Building Transparency into AI Projects. *Harvard Business Review*, 2022. június 20. <https://tinyurl.com/27xvyrst>

⁶³ EC Proposal 3.5. pont.

⁶⁴ DSA 65. preambulumbekzdés.

⁶⁵ Zódi Zsolt: Átláthatósági és indokolási követelmények az európai platformjogban. *In Medias Res*, 2023/2. 17. <https://doi.org/10.59851/imr.12.2.1>

⁶⁶ DSA 15. cikk.

⁶⁷ DSA 24. cikk.

⁶⁸ DSA 42. cikk.

⁶⁹ DSA 15. cikk (1) c).

⁷⁰ DSA 15. cikk (1) e).

⁷¹ DSA 42. cikk (2) c).

⁷² DSA 9. preambulumbekzdés.

szolgáltatók indokolási kötelezettsége mindazokra a korlátozásokra fennáll, amelyeket a felhasználó által megosztott tartalom (információ) jogellenessége vagy a szerződési feltételekkel való összeegyeztethetlensége miatt alkalmaznak, és fontos eleme, hogy a közvetítő szolgáltató a bejelentést követő passzivitását is köteles indokolni. A DSA a jogi döntésekre előírt felépítést, illetve logikát követi azzal a kivétellel, hogy az alkalmazott jogkövetkezményt is az indokolás tartalmazza a rendelkező rész helyett. Ezen túl kötelező elem a tényállás megállapítása, a jogalap vagy szerződéses feltétel megjelölése és külön pontban foglalnak helyet a döntés meghozatala során használt automatizált (mesterséges intelligencia vezérelt) eszközök alkalmazására vonatkozó információk. Ugyancsak a jogi döntések mintájára az indokolásnak tartalmaznia kell egy jogorvoslati klauzulát.

Ad (iv) Míg a DSA 17. cikkében előírt indokolási kötelezettség nagyrészt a szolgáltatók aktivitása esetén irányadó, a DSA 20. cikkében foglalt, legalább hat hónapig biztosítandó, elektronikus és ingyenes hatékony panaszmechanizmus⁷³ a döntések szélesebb körére vonatkozik, és abban az esetben is lehetőséget nyújt a panasz benyújtására, ha a 16. cikk szerinti bejelentés alapján a platform passzív marad.⁷⁴ A panasz eljárás szükségképpen személyzeti közrehatást igényel, és a szolgáltató indokolatlan késedelem nélkül köteles tájékoztatni a panaszost indokolt döntéséről és a peren kívüli vitarendezés lehetőségeiről vagy egyéb jogorvoslati lehetőségről. Az „indokolatlan késedelem nélkül” kitétel azonban legalább annyira tágan megfogalmazott követelmény, mint az Eker. irányelvben foglalt haladéktalan eltávolítási kötelezettség, és konkretizálás hiányában várhatóan számos bizonytalanság forrása lesz.⁷⁵

A 17. cikkben írtak alapján jogosan merül fel a kérdés, hogy mennyiben tud érdemben panaszt tenni az olyan bejelentést tevő, akinek a bejelentése alapján a szolgáltató nem tesz intézkedést, azonban a 17. cikk alapján indokolási kötelezettsége sem volt a döntés kapcsán. A korlátozás és a közvetítő szolgáltató passzivitása⁷⁶ esetére előírt eltérő indokolási kötelezettség azonban azzal az aszimmetrikus következménnyel járhat, hogy a felhasználó nem ugyanolyan esélyekkel tud jogorvoslattal élni a szolgáltató passzivitása (tartózkodása, nem tevése) esetén, mint az a felhasználó, akinek tartalmát a szolgáltató indokolt döntéssel távolította el.⁷⁷ A kötelezettségek összességét megvizsgálva azonban láthatjuk, hogy a szolgáltató köteles a körülményekre és a tényállásra kiterjedő indokolást nyújtani a belső panaszkezelési mechanizmus keretében eldöntött panasz kapcsán, így komplementálva a teljes indokolási kötelezettség hiányát passzivitás esetén (hiszen a peren kívüli vitarendezés igénybevételének előfeltétele a panasz-

⁷³ DSA 20. cikk (1).

⁷⁴ DSA 17. cikk.

⁷⁵ Részletesen ld. Gosztonyi Gergely: A közösségi média felelősségi kérdéseinek korai szabályozása az Amerikai Egyesült Államokban és az Európai Unióban. In: Mezey Barna (szerk.): *Kölcsönhatások. Európa és Magyarország a jogtörténelem sodrásában*. Budapest, Gondolat, 2021. 111–119.

⁷⁶ Pietro Ortolani: If You Build It, They Will Come. In: Joris van Hoboken et al. (szerk.): *Putting the DSA into Practice. Enforcement, Access to Justice and Global Implications*. Berlin, Verfassungsblog GmbH, 2023. 151–166. <http://www.doi.org/10.17176/20230208-093135-0>

⁷⁷ DSA 16. cikk (5).

mechanizmus kimerítése),⁷⁸ ezáltal lehetővé téve a felhasználónak, hogy a 21. cikk szerinti peren kívüli vitarendezési fórum előtt egy indokolt döntést tudjon vitatni.⁷⁹

Ad (v) A peren kívüli vitarendezésnek azért is van kiemelkedő jelentősége, mert korábban – ide nem értve a Facebook (Meta) által felállított Ellenőrző Bizottság közel négy évnyi tevékenységét – a felhasználók és a közösségi média platformok közötti vitarendezés egyetlen módja az állami (és esetlegesen választott-) bíróságon keresztül történt. 2024. február 17-től, a DSA teljeskörű alkalmazását követően a felhasználók peren kívüli vitarendezésre jogosultak, amelyet a rendeleti forma miatt minden tagállamnak biztosítani kell úgy, hogy a költségek szignifikáns részét a közösségi médiaszolgáltatók kötelesek viselni.⁸⁰ A rendelet által felvázolt peren kívüli vitarendezés mechanizmusa minden mikro- és kisvállalkozásnál nagyobb online platformot üzemeltető szolgáltatóra vonatkozik,⁸¹ aki/amely az EU-ban található, vagy ott van a letelepedési helye, függetlenül az említett közvetítő szolgáltatást biztosító szolgáltatók letelepedési helyétől.⁸²

Ugyan a DSA 21. cikkében foglalt testület előtti eljárás eredményeként meghozott döntések – az alávetés kivételével – kötelező erővel nem bírnak, lehetővé teszik, hogy releváns szakértelemmel rendelkező tagokból álló bizottság idő- és költséghatékony módon, a bírósági szervezetrendszeren kívül dönthessen olyan vitás kérdésekben, amelyeket az online platform felhasználóinak a szolgáltatóval közvetlenül nem sikerül rendezniük.⁸³ Ezen eljárások a bírósághoz fordulás jogának sérelme nélkül vehetők igénybe, de mindenképpen azzal szembeni előnyük, hogy egy szakértői testület hoz döntést, míg a bíróság legfeljebb csak bizonyos szakkérdésben támaszkodhat szakértői véleményekre.⁸⁴

3.2. Az MI rendelet és a tartalommoderálás, illetve az átláthatóság

Az MI rendelet jelentős változásokat hoz az online platformok mesterséges intelligencia tartalommoderálásban való felhasználásának módjában, ami lehetőségeket és kihívásokat egyaránt rejt magában. Az MI rendelet egy kockázatalapú megközelítést alkalmaz, és az AI-rendszereket a gyakorlatban négy kategóriába sorolja: elfogadha-

⁷⁸ Lendvai Gergely Ferenc: A Facebook Ellenőrző Bizottság működése és bíraskodása a gyűlöletbeszéd kontextusában. *In Medias Res*, 2024/1. 195–221. <https://doi.org/10.59851/imr.13.1.11>

⁷⁹ Husovec i. m. 24.

⁸⁰ Hannah Ruschemeier et al.: Brave New World: Out-Of-Court Dispute Settlement Bodies and the Struggle to Adjudicate Platforms in Europe. *Verfassungsblog*, 2024. szeptember 10. <https://tinyurl.com/yt9zn7pb> <http://www.doi.org/10.59704/46b8611eb2d96a84>

⁸¹ Gyetván Dorina: A Digital Services Act és a Facebook Oversight Board szerepe a jogorvoslat biztosításában a közösségi média vonatkozásában. *In Medias Res*, 2023/2. 190–208. <https://doi.org/10.59851/imr.12.2.9>

⁸² DSA 21. cikk.

⁸³ DSA 21. cikk (2).

⁸⁴ Gyetván Dorina – Gosztonyi Gergely: Garanciális elemek az online platformok véleménynyilvánítást korlátozó moderálási gyakorlatával szemben a digitális szolgáltatásokról szóló rendelet (DSA) előtt és után. *Magyar Jog*, 2025/4. 226–236. <https://doi.org/10.59851/mj.72.04.3>

tatlan, nagy, rendszerszintű és minimális kockázatú rendszerek.⁸⁵ Az online platformok által használt tartalommoderációs eszközök általában a magas vagy a korlátozott kockázatú rendszerek kategóriájába tartoznak, attól függően, hogy milyen hatással lehetnek az alapvető jogokra, például a véleménynyilvánítás szabadságára és a megkülönböztetésmentességre.

Az MI rendelet értelmében tiltott az olyan mesterséges intelligencia rendszerek használata, amelyek elfogadhatatlan kockázatot jelentenek az alapvető jogokra nézve. Ide tartoznak az olyan mesterséges intelligenciával működő rendszerek, amelyek manipulálják az emberi viselkedést, vagy olyan módon használják ki a kiszolgáltatott személyeket, ami kárt okozhat nekik.⁸⁶ A tartalommoderálással összefüggésben ebbe a kategóriába tartoznának azok az AI-rendszerek, amelyek szándékosan vagy szisztematikusan cenzúrázzák a törvényes tartalmakat a közvélemény manipulálása érdekében.⁸⁷

A nagy kockázatú mesterséges intelligencia rendszerekre a legszigorúbb szabályok vonatkoznak.⁸⁸ Az online óriásplatformok, például a Facebook, a YouTube és az X/ Twitter által használt tartalommoderációs rendszerek a felhasználók alapvető jogaira gyakorolt potenciális hatásuk miatt a magas kockázatúak közé sorolhatók. Ezek a platformok gyakran használnak mesterséges intelligenciát a közösségi normákat sértő tartalmak, például a gyűlöletbeszéd, a félretájékoztatás és a szélsőséges tartalmak automatikus felismerésére és eltávolítására. Tekintettel e funkciók érzékenységére, az EU különleges kötelezettségeket ró a magas kockázatú mesterséges intelligencia rendszerekre, különösen az átláthatóság, a felügyelet és a pontosság tekintetében.⁸⁹

A felhasználók jogait kevésbé befolyásoló tartalommoderációs eszközök a rendszer-szintű és a minimális kockázatú rendszerek közé tartoznak.⁹⁰ Ezek közé tartozhatnak olyan alapvető automatikus szűrőrendszerek, amelyek például blokkolják a spameket vagy észlelnek bizonyos szerzői jogi jogsértéseket. A korlátozott kockázatú rendszerek esetében a követelmények enyhébbek, de továbbra is tartalmazznak átláthatósági kötelezettségeket. A minimális kockázatú mesterséges intelligenciára, például az ajánlórendszerekben használt általános célú algoritmusokra vagy a korlátozott moderálási funkcióval rendelkező nyelvi szűrőkre még kevesebb szabályozási kötelezettség vonatkozik.⁹¹

⁸⁵ Martin Ebers: Truly Risk-Based Regulation of Artificial Intelligence. How to Implement the EU's AI Act. *SSRN*, 2024. június 19., 7–11. <http://dx.doi.org/10.2139/ssrn.4870387>

⁸⁶ MI rendelet 5. cikk.

⁸⁷ Michael Veale – Frederik Zuiderveen Borgesius: Demystifying the Draft EU Artificial Intelligence Act. Analysing the good, the bad, and the unclear elements of the proposed approach. *Computer Law Review International*, Vol. 22., No. 4. (2021) 97–112. <https://doi.org/10.9785/cr-2021-220402>

⁸⁸ MI rendelet 6–27. cikkek.

⁸⁹ Gerhard Wagner: Liability for Artificial Intelligence. A Proposal of the European Parliament. *Working Paper des Forschungsinstituts für Recht und digitale Transformation*, 2021/9. 1–34. <https://doi.org/10.2139/ssrn.3886294>

⁹⁰ MI rendelet 51–55. cikkek.

⁹¹ Luciano Floridi: The European Legislation on AI: a Brief Analysis of its Philosophical Approach. *Philosophy & Technology*, Vol. 34., No. 1. (2021) 215–222. <https://doi.org/10.1007/s13347-021-00460-9>

Az MI rendelet több cikke közvetlenül foglalkozik a mesterséges intelligenciát tartalmazó moderálására használó platformok felelősségével, az átláthatóságra, az adatkezelésre, az elszámoltathatóságra és az emberi felügyeletre összpontosítva.⁹²

A tartalom moderálásához használt mesterséges intelligencia rendszerek tisztességes és elfogulatlan működésének biztosításához elengedhetetlen a megfelelő adatkezelés.⁹³ Az MI rendelet 10. cikke előírja, hogy a magas kockázatú mesterséges intelligencia rendszerek képzéséhez felhasznált adatoknak jó minőségűnek, elfogultságtól mentesnek és az összes releváns csoportot reprezentálnak kell lenniük. Ez különösen fontos a tartalommoderálásban, ahol az elfogult képzési adatok bizonyos csoportokkal szembeni tisztességtelen bánásmódot vagy a marginalizált közösségek túlságosan szigorú ellenőrzését eredményezhetik.⁹⁴ Ha például a gyűlöletbeszéd megjelölésére használt mesterséges intelligencia rendszert olyan adatokon képzik ki, amelyekben túlnyomórészt bizonyos dialektusok vagy nyelvek szerepelnek, akkor az aránytalanul nagy mértékben megjelölheti az ezekből a csoportokból származó tartalmakat, ami aggályokat vethet fel a diszkriminációval és a szólásszabadsággal kapcsolatban.

Az automatizált döntésekre való túlzott hagyatkozás megelőzése érdekében az MI rendelet 14. cikke előírja, hogy a nagy kockázatot jelentő mesterséges intelligencia rendszereket emberi felügyeletnek kell alávetni.⁹⁵ A tartalom moderálásával összefüggésben ez azt jelenti, hogy bár a mesterséges intelligencia automatizálhatja a káros tartalmak felismerését, a végső döntések meghozatalában – különösen a felhasználók alapvető jogait érintő döntésekben – emberi moderátoroknak kell részt venniük. Ez az ember által alkalmazott megközelítés elengedhetetlen a mesterséges intelligencia korlátainak kezeléséhez, például ahhoz, hogy nem képes teljes mértékben felfogni vagy helyén kezelni a kontextust. A tartalommoderálás esetében ez azt jelenti, hogy az AI-eszközök megbízhatóan felismerik a káros tartalmakat, miközben minimalizálják a hamis pozitív eseteket, amikor a jogszerű tartalmakat tévesen jelölik vagy távolítják el, és a hamis negatív eseteket, amikor a káros tartalmakat nem veszik észre.⁹⁶ Az emberi felügyeletre vonatkozó rendelkezések megerősítik az automatizált döntések és az emberi ítélőképesség egyensúlyának fontosságát. Míg a mesterséges intelligencia hatékonyan képes nagy mennyiségű tartalom moderálására, az emberi moderátorok továbbra is nélkülözhetetlenek az összetett, kontextuális megértést igénylő ügyek kezeléséhez.

⁹² Solymár Károly Balázs: A fogalomhasználat egyes kérdései a mesterséges intelligenciáról szóló európai rendelettervezetben. In: Glavanits Judit (szerk.): *Fogyasztóbarát mesterséges intelligencia – a velünk élő AI egyes aktuális kérdései*. Győr, Universitas-Győr Nonprofit Kft., 2023. 9–27.

⁹³ Theodore S. Boone: The challenge of defining artificial intelligence in the EU AI Act. *Journal of Data Protection and Privacy*, Vol. 6., No. 2. (2023) 180–195. <https://doi.org/10.69554/QHAY8067>

⁹⁴ Reuben Binns: Fairness in Machine Learning: Lessons from Political Philosophy. *Proceedings of Machine Learning Research*, Vol. 10., No. 81. (2018) 149–159.

⁹⁵ Tóth András: Az Európai Unió Mesterséges Intelligencia Törvényéről. *Gazdaság és Jog*, 2024/5–6. 3–11.

⁹⁶ Tarleton Gillespie: *Custodians of the Internet. Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media*. New Haven, Yale University Press, 2018. 99. <https://doi.org/10.12987/9780300235029>

Az MI rendelet bizonyos, tartalommoderálási kérdésekkel is összefüggésbe hozhatóan, a kockázatok csökkentése érdekében viszonylag hangsúlyosan tartalmaz átláthatósági követelményeket, meghatározásokat a DSA-hoz hasonlóan. A fogalom tisztázása érdekében például a 27. preambulumbekkezdés leszögezi, hogy az MI rendelet keretei között az átláthatóság azt jelenti, hogy az MI-rendszereket olyan módon kell fejleszteni és használni, amely lehetővé teszi a megfelelő nyomonkövethetőséget és megmagyarázhatóságot, miközben tudatosítja az emberekben, hogy MI-rendszerrel kommunikálnak vagy lépnek kapcsolatba, valamint megfelelően tájékoztatja az alkalmazókat az MI-rendszer képességeiről és korlátairól, az érintett személyeket pedig jogaikról.⁹⁷

Erre építve az MI rendelet 50. cikke a különböző érdekelt jogainak és érdekeinek egyensúlyban tartására, az információs aszimmetria mérséklésére törekedve tartalmaz egyes olyan MI-rendszerekre vonatkozóan átláthatósági kötelezettségeket, amelyek (a) emberekkel érintkeznek, (b) az érzelmek észlelésére vagy a biometrikus adatokon alapuló (szociális) kategóriákkal való kapcsolat meghatározására szolgálnak, vagy (c) tartalmakat hoznak létre vagy manipulálnak.⁹⁸ Ha a emberek MI-rendszerrel érintkeznek, vagy érzéseiket vagy jellemzőiket automatizált eszközökkel ismerik fel, erről őket tájékoztatni kell. Ha az MI-rendszert olyan képek, audio- vagy videotartalmak létrehozására vagy manipulálására használják, amelyek érzékelhetően hasonlítanak hiteles tartalmakra, kötelezővé kell tenni annak közlését, hogy a tartalom automatizált módon jött létre, eltekintve a jogszerű célokat szolgáló kivételektől. Ez lehetővé teszi a személyek számára, hogy megalapozott döntéseket hozzanak, vagy visszalépjenek adott helyzetből.⁹⁹

4. A mesterséges intelligencia a tartalommoderálás gyakorlatában

A mesterséges intelligencia központi eszközzé vált az online platformok tartalommoderációjában, ahol a felhasználók által generált hatalmas mennyiségű tartalom kezelése kritikus kihívást jelent. Az olyan platformok, mint a Facebook, az X/ Twitter, a YouTube és az Instagram naponta hatalmas mennyiségű posztot, képet, videót és kommentet dolgoznak fel, és az AI hatékony módot kínál a közösségi irányelveket sértő tartalmak ellenőrzésének, megjelölésének és eltávolításának automatizálására. Bár a mesterséges intelligencia nagymértékben javíthatja a tartalom moderálását, nem mentes a kihívásoktól. Alkalmazása sikeres és hibás tartalomosztályozást egyaránt eredményezett, ami számtalan vitához vezetett a pontosság, a tisztességesség és az etikai vonatkozások körül. Jelen fejezetben vizsgáljuk a tartalom moderálásának hatékonyságát, korlátait és azokat a konkrét eseteket, amikor a mesterséges intelligencia jól vagy rosszul teljesített.

A mesterséges intelligencia a tartalom moderálásában általában a gépi tanulási algoritmusok, a természetes nyelvi feldolgozás és a számítógépes látás kombinációjával működik. Ezek a technológiák teszik lehetővé a mesterséges intelligencia számára,

⁹⁷ MI rendelet 27. preambulumbekkezdés.

⁹⁸ MI rendelet 50. cikk.

⁹⁹ EC Proposal 5.2.4.

hogy nagy mennyiségű adatot gyorsan megvizsgáljon és elemezzen, és olyan mintákat vagy jeleket azonosítson, amelyek a közösségi irányelvek megsértésére utalhatnak.

Az algoritmusokat nagy adathalmazokon tanítják be, amelyek az elfogadható és elfogadhatatlan tartalom címkézett példáit tartalmazzák. A mesterséges intelligencia ezután e tanult minták alapján megjósolja, hogy az új tartalom megfelel-e a közösségi irányelveknek.¹⁰⁰ Az ilyen modellek folyamatosan fejlődnek, ahogy új adatokkal találkoznak, lehetővé téve a tartalomfelismerés dinamikus fejlesztését. A természetes nyelvi feldolgozás (*Natural Language Processing*, a továbbiakban: NLP) elengedhetetlen a szövegalapú tartalom elemzéséhez és a nem megfelelő kifejezések felismeréséhez. Az NLP modellek néha képesek felismerni egy szó vagy kifejezés kontextusát, megkülönböztetve a jóindulatú és a káros felhasználásokat.¹⁰¹ Az X/Twitter például NLP-t használ a sértő nyelvezetet vagy gyűlöletbeszédet tartalmazó tweetek megjelölésére. A képek és videók moderálásához a számítógépes látás algoritmusai elemzik a vizuális tartalmat, hogy felismerjék a nem megfelelő képi anyagokat, például a meztelenséget, a grafikus erőszakot vagy a terrorista propagandát.¹⁰²

A mesterséges intelligencia számos esetben nagyon hatékony a káros tartalmak azonosításában. A hatalmas mennyiségű tartalom gyors feldolgozásának képessége elengedhetetlen a több milliárd felhasználóval rendelkező platformok számára. A Facebook például sajtóhírek szerint jelentős összegeket fektetett be a mesterséges intelligenciába a terrorizmussal kapcsolatos tartalmak felderítése és eltávolítása érdekében. A vállalat 2024-ben arról számolt be, hogy a terrorizmussal kapcsolatos posztok több mint 99%-át az AI megjelölte, mielőtt még azokat a felhasználók jelentették volna.¹⁰³ A COVID-19 világvjárvány idején a YouTube mesterséges intelligenciát használt az egészségügyi félretájékoztatók terjedése elleni küzdelemhez, és a mesterséges intelligencia segítségével azonosították az irányelvek megsértése miatt megjelölt videók 94%-át.¹⁰⁴ A Twitter mesterséges intelligenciája 2024-ben a jogsértő felhasználói fiókok több mint 99%-át proaktívan kezelte anélkül, hogy felhasználói jelentésekre támaszkodott volna.¹⁰⁵

E sikerek ellenére a mesterséges intelligencia alapú tartalommoderáció még mindig jelentős korlátokkal küzd, ami nagy visszhangot kiváltó kudarcokhoz vezetett. Ezek gyakran abból fakadnak, hogy a mesterséges intelligencia nem képes megérteni a

¹⁰⁰ Robert Gorwa – Reuben Binns – Christian Katzenbach: Algorithmic content moderation: Technical and political challenges in the automation of platform governance. *Big Data & Society*, Vol. 7., No. 1. (2020) <https://doi.org/10.1177/2053951719897945>

¹⁰¹ Anna Schmidt – Michael Wiegand: A Survey on Hate Speech Detection using Natural Language Processing. In: Lun-Wei Ku – Cheng-Te Li (szerk.): *Proceedings of the Fifth International Workshop on Natural Language Processing for Social Media*. Valencia, Association for Computational Linguistics, 2017. 1–10. <https://doi.org/10.18653/v1/W17-1101>

¹⁰² Theo Araujo et al.: In AI we trust? Perceptions about automated decision-making by artificial intelligence. *AI & Society*, Vol. 35., No. 3. (2020) 611–623. <https://doi.org/10.1007/s00146-019-00931-w>

¹⁰³ Facebook: *Community Standards Enforcement Report. Q2 2024 report*. <https://tinyurl.com/22k4j7kp>

¹⁰⁴ Lambodara Parabhoi et al.: YouTube as a source of information during the Covid-19 pandemic: a content analysis of YouTube videos published during January to March 2020. *BMC Medical Informatics and Decision Making*, Vol. 21., No. 1. (2021) 255–265. <https://doi.org/10.1186/s12911-021-01613-8>

¹⁰⁵ X: *Global Transparency Report, H1 2024*. <https://tinyurl.com/2p86w9bx>

tartalom teljes kontextusát, a kulturális árnyalatokat és a nyelv bonyolultságát. Bár e rendszerek folyamatosan fejlődnek, a hibás tartalommoderáció számos reputációs problémát tud okozni a vállalatoknak is. 2016-ban a Facebook mesterséges intelligenciája helytelenül eltávolította a vietnami háborúban készült ikonikus Napalm Girl fotót, arra hivatkozva, hogy meztelensége sérti a tartalmi irányelveket.¹⁰⁶ Az MI nem ismerte fel a kép történelmi jelentőségét, és kizárólag a meztelenség alapján nem megfelelő tartalomként kezelte. 2019-ben az égő párizsi Notre Dame katedrális képeit távolította el az MI, mivel – helytelenül – a 2001-es amerikai terrortámadás képeinek értelmezte azokat.¹⁰⁷ A 2020-as COVID-19 világjárvány idején a YouTube szintén nagymértékben támaszkodott a mesterséges intelligenciára a tartalom moderálásában, mivel az emberi moderátorok otthonról dolgoztak. Ez a hamis pozitív eredmények jelentős növekedéséhez vezetett, amikor a COVID-19-ről szóló oktatási és hírekkel kapcsolatos tartalmakat tévesen félretájékoztatásként távolították el. De a satirikus tartalom is gyakran kihívást jelent a mesterséges intelligencia számára, ugyanis a satírárt – annak a normál beszédőtől eltérő kontextusáról és hangneméről szinte tudomást sem véve – gyakran azonosítja a mesterséges intelligencia jogsértő tartalomként.¹⁰⁸

A mesterséges intelligenciának a tartalom moderálásával kapcsolatos nehézségei nagyrészt abból erednek, hogy a betáplált adatmodellek hatalmas mennyisége ellenére sem képes teljes mértékben megérteni a kontextust, az eltérő kultúrákat és a nyelvi különbségeket. Így például fontos lenne annak ismerete, hogy a nyelv folyamatosan fejlődik, és gyorsan jelennek meg új szleng- és köznyelvi kifejezések, különösen a közösségi médiában. Az elavult adathalmazokon kiképzett mesterséges intelligencia-modellek nem feltétlenül képesek megragadni az új kifejezéseket, ami pontatlan moderáláshoz vezethet. Ráadásul a különböző kultúrákban a nyelvezet jelentősen eltér, ami megnehezíti, hogy a globális platformok minden országban egyformán és egységesen alkalmazható moderációs rendszereket fejlesszenek ki. Nem szabad elfelejteni, hogy a mesterséges intelligencia-modellek csak annyira jók, mint az adatok, amelyeken kiképzik őket. Ha az adatok elfogultak vagy hiányosak, a végeredmény is elfogult eredményeket fog produkálni.

5. Összefog(lal)ás? A mesterséges intelligencia és az emberi moderálás kapcsolatának jövője

A mesterséges intelligencia korlátai miatt sok platform hibrid megközelítést alkalmaz a tartalom moderálására, amely a mesterséges intelligencia sebességét emberi felügyelettel kombinálja. Az emberi moderátorok felülvizsgálják a megjelölt tartalmakat, hogy biztosítsák a kontextuális és nyelvi árnyalatok, illetve a kulturális különbségek figyelembevételét. Ez az együttműködő megközelítés lehetővé teszi a platformok számára,

¹⁰⁶ Gillespie i. m. 1–4.

¹⁰⁷ Tuesday briefing: YouTube's fake news detection tool flagged the Notre Dame fire with 9/11 facts. *Wired*, 2019. április 16., <https://tinyurl.com/5547uudk>

¹⁰⁸ Tom Lynn – Jessica Bancroft: The use of algorithms in the content moderation process. *gov.uk*, 2021. augusztus 5. <https://tinyurl.com/hryn9uy8>

hogy egyensúlyba hozzák a mesterséges intelligencia erősségeit az árnyalt döntéshozatal emberi képességével. A mesterséges intelligencia fejlődésével a tartalommoderációs rendszerek valószínűleg egyre kifinomultabbá válnak. A mélytanulás, a kontextuális elemzés és a mesterséges intelligencia kulturális különbségek megértésére való képességének javulása növelheti a moderációs rendszerek pontosságát.

Az EU mesterséges intelligencia rendelete és a DSA fontos szerepet fognak játszani az online platformokon a mesterséges intelligencia által vezérelt tartalommoderáció jövőjének alakításában. A rendeletek szigorú követelményeket írnak elő a mesterséges intelligenciával működő rendszerekkel szemben, és azt kívánják biztosítani, hogy a tartalommoderációs eszközök átláthatóak, tisztességesek és elszámoltathatók legyenek. Ezzel lehetőséget kínálnak arra, hogy a tartalommoderálás minőségének javítása és az alapvető jogok védelme a digitális korban is biztosított legyen.

Összefoglalva, bár a mesterséges intelligencia létfontosságú szerepet játszik az online platformok tartalmi moderálásában, annak használata és használhatósága nem korlátlan.¹⁰⁹ A mesterséges intelligencia és az emberi moderálás közötti egyensúly – ahogy az online platformok alkalmazkodnak majd az új szabályozásokhoz – továbbra is fejlődni fog, különösen azáltal, hogy a platformok is rá lesznek kényszerítve arra, hogy emberi közreműködésen keresztül enyhítsék a felmerülő technológia-indukált problémákat. Az emberi felügyelet a közeljövőben továbbra is kulcsfontosságú marad a mesterséges intelligencia által még meg nem érthető finomságok és összetettségek kezeléséhez.

¹⁰⁹ Content moderation in a new era for AI and automation. *Oversight Board*, 2024. szeptember 17. <https://tinyurl.com/yc5b4nbs>

ÁTLÁTHATÓSÁGOT BIZTOSÍTÓ SZABÁLYOK ÉS INTÉZKEDÉSEK AZ MI TARTALOMMODERÁLÁSI TEVÉKENYSÉGE SORÁN – KIEMELT FIGYELEMMEL A DSA KÖVETELMÉNYRENDSZERÉRE*

Hohmann Balázs**

1. Bevezetés, problémameghatározás

A mesterséges intelligenciára épülő (MI) technológia megoldások meghatározó részeivé váltak a mindennapjainknak, s így alapvető hatást fejtenek ki a közösségi platformok, online terek tartalommoderálási tevékenységére is.¹ A nagy mennyiségű adat feldolgozásán, valós időben történő döntéselőkészítésen és döntéshozatalon alapuló rendszerek, melyek az emberi beavatkozás minimalizálására is szolgálnak, alapvető mozgatórugói lettek minden, a platformjog által szabályozott szolgáltatásnak.² A változás hatással van a közösségi érintkezés és kommunikáció módszerére is,³ olyannyira, hogy a társadalmi interakció átalakulásával az ezen a téren mutatkozó technológiai fejlődés és az abból eredő hatások már egyértelműen nemcsak eszközként jelennek meg, hanem a felhasználói döntéseket, a platform jogi megfelelőségét alapvetően befolyásoló tényezővé váltak, a társadalom szinte minden folyamatát tekintve.⁴

* Készült a Nemzeti Média- és Hírközlési Hatóság támogatásával. A tanulmányba foglalt kutatási eredmények létrehozása részben a Kulturális és Innovációs Minisztérium Nemzeti Kutatási Fejlesztési és Innovációs Alapból nyújtott támogatásával, az NKKP STARTING_24 pályázati program finanszírozásában valósult meg, a STARTING 150399 számú projekt keretében.

** Egyetemi adjunktus, Pécsi Tudományegyetem, Állam- és Jogtudományi Kar.
ORCID: <https://orcid.org/0000-0003-4252-1923>

¹ Pintér Melinda: Az online platformok diskurzusának moderációja és biztonsági aggályai. *Scientia et Securitas*, Vol. 3., No. 1. (2022) 45. <https://doi.org/10.1556/112.2022.00089>

² Tóth András: Az online platformok európai szabályozása. *In Medias Res*, Vol. 11., No. 2. (2022) 78–81. <https://doi.org/10.59851/imr.11.2.6>

³ Papp János Tamás: *A közösségi média platformok szabályozása a demokratikus nyilvánosság védelmében*. Budapest, Wolters Kluwer, 2022. 7–8.

⁴ Hadi Rhonda – Shiri Melumad – Eric S. Park: The Metaverse: A new digital frontier for consumer behavior. *Journal of Consumer Psychology*, Vol. 34., No. 1. (2024) 142–150. <https://doi.org/10.1002/>



A tanulmány ezzel kapcsolatban egy összetett alapvetésből indul ki: a tartalommoderálási tevékenység napjainkra nem megvalósítható automatizált megoldások bevonása nélkül. Egyrészt a vezető közösségi média platformokon keresztül olyan mennyiségű moderálható tartalom halad át,⁵ amelyet emberi erővel, teljes terjedelmében gazdaságtalan lenne megvizsgálni. Másrészt a közösségi média platformok és egyéb kapcsolódó platformszolgáltatások mára olyan jelentős szerepet töltenek be a társadalmi közvélemény formálásában,⁶ amely alapján megengedhetetlen az is, hogy platform részéről ne történjen semmilyen érdemi moderációs tevékenység, valamint az sem, hogy ennek során, a gépesítés hiányában, időbeli késedelem, vagy egyedi emberi döntésekből kialakuló eltérő következmények alakuljanak ki azonos tartalmakat feltételezve, mert az korlátozza a véleményppluralizmust a demokratikus közvélemény-formálódás folyamatában.⁷ Azt is figyelembe kell venni, hogy a platformok üzleti modelljeinek középpontjában az algoritmusokkal támogatott tartalomajánlás- és moderálás áll.⁸ Az elmúlt években éppen ezért a platformok működésének egyik legfontosabb vizsgált területe a tartalommoderációs gyakorlatok átláthatósága és jogi megfelelőségének biztosítása lett.⁹ A kiskorúak védelme érdekében, a jogsértő deepfake-tartalmak, a gyűlöletbeszéd vagy a terrorizmus elleni fellépés vagy egyéb jogsértő tartalmak tekintetében alkalmazott moderációs megoldások azonban folyamatosan a „határon” táncolnak. Ha nem megfelelő szűrést biztosítanak, jogellenes tartalmak kerülnek nagy tömegek elé, míg a túl szigorú szabályok szerint haladó moderálás könnyen korlátozhatja a véleménynyilvánítás szabadságát, más alapvető jogosultságok mellett.¹⁰ A „tűlszűrés” jelensége is megjelenhet, amely a jogszerű, de érzékeny vagy vitatott témákat is kiszűri, korlátozva a vélemények szabad megjelenését.

jcpy.1356

- ⁵ Leonardo Madio – Martin Quinn: Content moderation and advertising in social media platforms. *Journal of Economics & Management Strategy*, Early view (2024) 345. <https://doi.org/10.1111/jems.12602>; Young K. Greyson: How much is too much: the difficulties of social media content moderation. *Information & Communications Technology Law*, Vol. 31., No. 1. (2022) 5. <https://doi.org/10.1080/13600834.2021.1905593>
- ⁶ Teleki Lóránt – Gönczöl Tünde: A digitális kapuőrök az ex ante szabályozás, a versenyjog és az adatvédelem keretében. *Verseny és Szabályozás*, 2022/1. 243.
- ⁷ Gosztonyi Gergely: A tartalomszabályozás emberi és technikai vonatkozásai. *Erdélyi Jogélet*, Vol. 4., No. 4 (2021) 12–14.
- ⁸ Rob Frieden: The internet of platforms and two-sided markets: Implications for competition and consumers. *Villanova Law Review*, Vol. 63., No. 2. (2018) 269. <https://doi.org/10.2139/ssrn.2935292>
- ⁹ Török Bernát: A szólásszabadság a közösségi platformokon és a Digital Service Act. In: Török Bernát – Zódi Zsolt (szerk.): *Az internetes platformok kora*. Budapest, Ludovika, 2022. 195–200.; Jens Brunk – Jana Mattern – Dennis M. Riehle: Effect of transparency and trust on acceptance of automatic online comment moderation systems. In: Jörg Becker et. al. (szerk.): *2019 IEEE 21st Conference on Business Informatics (CBI) I*. Moscow, IEEE, 2019. 429–433. <https://doi.org/10.1109/CBI.2019.00056>; Nicolas P. Suzor – et al.: What do we mean when we talk about transparency? Toward meaningful transparency in commercial content moderation. *International Journal of Communication*, Vol. 13., No. 1. (2019) 18.
- ¹⁰ Poppett Ryan: Is content moderation a concern that is best left to platforms? *Ethica*, Vol. 2., No.1. (2022) 73. Ld. pl. gyakorlati esettanulmányon keresztül: Heather Tillewein – Keely Mohon-Doyle – Destiny Cox: A Critical Discourse Analysis of Sexual Violence Survivors and Censorship on the Social Media Platform TikTok. *Archives of Sexual Behavior*, Vol. 53., No.1. (2024) 1–10. <https://doi.org/10.1007/s10508-024-02987-2>

Az algoritmusok működésével kapcsolatos átláthatóság hiánya kockázatokat jelent, amelyet sokáig nem is sikerült a felszínre hozni szabályozás segítségével. Az algoritmusok ugyanis sokáig az üzemeltető vállalkozások védett ismereteként, üzleti titkaként kerültek kezelésre,¹¹ ami megnehezítette a szolgáltatók ellenőrzését és felelősségre vonását. Az algoritmusok működésével kapcsolatos torzítások, például a politikai nézetek, társadalmi csoportok mentén vagy bizonyos felhasználók hátrányos megkülönböztetésével súlyosan sérthetik az emberi jogok érvényesülését.¹²

A jelen tanulmányban közelebbről megvizsgált digitális szolgáltatásokról szóló rendelet (továbbiakban: DSA),¹³ különös jelentőséggel bír az MI által támogatott tartalommoderálási rendszerek szempontjából is, mivel explicit előírásokat fogalmaz meg a platformszolgáltatások és ezek körében az algoritmus-vezérelt folyamatok átláthatóságára, az érintettek tájékoztatására és az alapjogok védelmére vonatkozóan. A szabályozás célja a már említett kockázatok minimalizálása, miközben biztosítja a digitális piac zavartalan működését. A rendelet rétegzett, kockázatalapú szabályozási megközelítést alkalmaz, amelynek célja, hogy a kockázattal kapcsolatos átláthatósági kötelezettségek terheljék a szolgáltatókat a szolgáltatáshoz kapcsolódó kockázatkezelési intézkedések mellett. A DSA keretrendszere számos olyan előremutató, kulcsfontosságú elemet tartalmaz, amelyek az átláthatóságot erősítik az MI-tartalommoderálási tevékenységek során, reflektálva a korábbi kritikákra és jogalkalmazási problémákra. Ezek közé tartoznak az átfogó tájékoztatási kötelezettségek, az alkalmazott MI rendszerek és a platform működésére vonatkozó jelentések, az emberi felülvizsgálat lehetősége, a kockázatelemzések és az átláthatósági jelentések közzététele.

Az átláthatósági követelmények vonatkozásában a DSA az online óriásplatformok és a nagyon népszerű keresőprogramokra állapít meg fokozott kötelezettségeket, az átláthatósági szabályoknak való megfelelés kérdése azonban kisebb platformszolgáltatók esetében is felmerül. E szolgáltatók gyakran nem rendelkeznek azokkal az erőforrásokkal, amelyek szükségesek az előírások maradéktalan teljesítéséhez, ami gátolhatja az innovációt és versenyhátrányt okozhat számunkra a nagyobb platformokhoz képest.¹⁴

¹¹ Mauro Conti et al.: Revealing The Secret Power: How Algorithms Can Influence Content Visibility on Social Media. *arXiv preprint arXiv:2410.17390* (2024) 2–3. <https://doi.org/10.48550/arXiv.2410.17390>; Edouard Pignot: Who is pulling the strings in the platform economy? Accounting for the dark and unexpected sides of algorithmic control. *Organization*, Vol. 30., No. 1. (2023) 150–151. <https://doi.org/10.1177/1350508420974523>

¹² Tatjana Jovanić: Platform Regulation: Ex Ante Tools and Strategies in the Regulatory Toolkit. In: Dušan V. Popović – Rainer Kulms (szerk.): *Repositioning Platforms in Digital Market Law*. Cham, Springer Nature Switzerland, 2024. 2–5. https://doi.org/10.1007/978-3-031-69678-7_1

¹³ 2022/2065 rendelet a digitális szolgáltatások egységes piacáról és a 2000/31/EK irányelv módosításáról HL L 277, 2022.10.27., 1–102. o.

¹⁴ Bence Kis Kelemen – Balázs Hohmann: Is There Anything New Under the Sun? A Glance at the Digital Services Act and the Digital Markets Act from the Perspective of Digitalisation in the EU. *Croatian Yearbook of European Law And Policy*, Vol. 19., No. 1. (2023) 247–248. <https://doi.org/10.3935/cyelp.19.2023.542>

Azt is célszerű jelen tanulmány szempontjából azonosítani, hogy az átláthatóság nem csupán jogi, hanem társadalmi elvárás is a platformszolgáltatókkal szemben.¹⁵ Az MI rendszerek döntéseinek és működési mechanizmusainak megértése alapvető fontosságú a felhasználók bizalmának elnyerése, valamint az alapjogok érvényesülésének biztosítása érdekében. E keretben a DSA olyan alapelveket és mechanizmusokat fogalmaz meg, amelyek segítik a platformszolgáltatókat abban, hogy működésük a társadalom széles köre számára is átlátható legyen.

A tanulmány célja, hogy átfogó képet nyújtson a DSA átláthatósági követelményeiről, különös tekintettel az MI-alapú tartalommoderálás sajátosságaira. A tanulmány arra a kérdésre keresi a választ, hogy milyen előrelépést jelent a DSA követelményrendszere az MI alapú tartalommoderálás átláthatóságára vonatkozóan? Milyen jogi és technikai kihívásokat támaszt a DSA által előírt követelmények teljesítése? Ezek a kérdések irányutatást nyújtanak a tanulmány tartalmi kereteinek meghatározásához, valamint hozzájárulnak a szabályozás és az MI alapú technológiák közötti kapcsolat alaposabb megértéséhez. A tanulmány doktrinális kutatási módszert alkalmaz, amely a vonatkozó jogszabályok és szakirodalom összehasonlító és kritikai elemzésén keresztül vizsgálja a fenti kérdéseket.

2. Út a DSA megalkotásáig – A szabályozást életre hívó okok

Jelen tanulmánynak nem célja, hogy részletesen elemezze a DSA megalkotásának folyamatát, azonban a DSA-ban a tartalommoderálás szempontjából fontos átláthatósági szabályok helyes értelmezéséhez és rendszerszemléletű megközelítése érdekében fontos foglalkozni a szabályozás előzményeivel is. Az európai szabályozás középpontjában a megelőző időszakban az elektronikus kereskedelmi irányelv szabályai álltak (2000/31/EK),¹⁶ amely a DSA megjelenését megelőzően (s azt követően részben hatályon kívül helyezett szabályokkal) szabályozta az információs társadalommal összefüggő szolgáltatások jogi kereteit. A szabályozás az Európai Unió digitális gazdaságának egyik első mérföldköve volt. Az irányelv célja a belső piac működésének elősegítése volt a digitális szolgáltatásokra vonatkozó jogi akadályok csökkentésével.¹⁷ Bár az irányelv alapvetően sikeresnek bizonyult az online gazdaság fejlődésének előmozdításában, számos jelentős hiányosságot hagyott maga után, amelyet a jogalkotó a modern platformjog és annak részeként a DSA megalkotásával igyekezett kiküszöbölni.¹⁸

¹⁵ Hohmann Balázs: A Digital Services Act és a Digital Markets Act termékekre és digitális szolgáltatásokra irányuló fogyasztói jogviszonyokat érintő rendelkezései. In *Medias Res*, Vol. 12., No. 2. (2023) 70–72. <https://doi.org/10.59851/imr.12.2.4>

¹⁶ 2000/31/EK irányelv a belső piacon az információs társadalommal összefüggő szolgáltatások, különösen az elektronikus kereskedelem, egyes jogi vonatkozásairól. HL L 178, 2000.07.17., 1–16. o. (a továbbiakban: e-kereskedelmi irányelv).

¹⁷ E-kereskedelmi irányelv, 1. cikk.

¹⁸ Klein Tamás: Az újmédia szabályozásának új iránya? – A Digital Services Act alapjogvédelmi mechanizmusa. In: Török Bernát – Zódi Zsolt (szerk.): *Digitalizálódó társadalom – Tanulmányok az új technológiák társadalmi-jogi hatásairól*. Budapest, Ludovika, 2023. 122–124.

Az e-kereskedelmi irányelv egyik legjelentősebb szabályozási megoldásának számított az előzetes engedélyezést kizáró elv (4. cikk).¹⁹ Ez az elv meghatározta a közvetítő szolgáltatók működésének kereteit. Az elv jelentősége a szolgáltatók terjeszkedésének elősegítésében, a verseny fellendítésében és az egységes digitális piac megalapozásában állt. Az elv alapján a tagállamok nem tehetik az ilyen szolgáltatások nyújtását előzetes engedélyeztetéshez kötötté.

Az elv alkalmazásának előnyei ellenére komoly joghatósági problémákat eredményezett, amelyek akadályozták a tagállami hatóságok hatékony felügyeleti tevékenységét a közvetítő szolgáltatások felett. Az irányelv e szabálya a „származási ország elvére” épült, amely kimondta, hogy egy adott digitális szolgáltatás felett az az ország gyakorol joghatóságot, ahol a szolgáltató letelepedett.²⁰ Ez azt jelentette, hogy például egy Írországban bejegyzett, de az EU számos tagállamában működő platform esetében az ír hatóságok voltak illetékesek a szabályozás és ellenőrzés terén. Az eltérő nemzeti szabályozások, valamint a joghatósági különbségek lehetőséget teremtettek a *forum shoppingra*, amely során a szolgáltatók előnyösebb szabályozási környezetet választottak letelepedési helyül. A probléma még súlyosabban jelentkezett akkor, amikor a szolgáltató nem telepedett le az Európai Unió területén, hanem szolgáltatásait harmadik országból nyújtotta az Európai Unió területén található jogalanyok számára. A szabályozás irányukba – akár tagállami hatóság, akár egyéb jogalany (felhasználó) kezdeményezte – szinte érvényesíthetetlen maradt, mert hiányzott az a mechanizmus, amely megteremtette volna az alapot a szolgáltatók felelősségre vonására, másrésről pedig felelősségre vonás esetén a jogkövetkezmények szinte érvényesíthetetlenek maradtak, mert a szolgáltatókat nem tette érdekeltté a szabályozás a szabályok betartására, komolyabb szankciók hiányában. A szabályozás irányelvi jellegéből adódóan ráadásul nem eredményezett egységes európai szabályozást, tagállami eltérések mutatkoztak, amelyek erősítették a fent már vázolt problémákat.

Az érintett szolgáltatások átláthatóságával kapcsolatos legfontosabb problémák éppen az említett jogalkalmazási problémákon alapultak. Előzetes engedélyezés hiányában a jogalkalmazó szervek nem voltak abban a helyzetben, hogy előre kiszűrjék azokat a problémákat, amelyek a tartalommoderáció területén mutatkoztak, a szolgáltatók pedig ennek hiányában maguk döntöttek arról, hogy mennyire teszik átláthatóvá a szolgáltatásukat. Kötelező szempontrendszer hiányában a szolgáltatók semmilyen módon nem voltak ösztönözve a tartalommoderációs tevékenység láthatóvá tételében, értelmezésében, amelyek javították volna a felhasználók átláthatóság-érzetét és egyúttal áttekinthetővé tették volna a szolgáltatás nyújtásának ezirányú feltételeit is.

A szabályozás másik fókuszponti problémája a közvetítő szolgáltatók fajtái és a felelősségi szabályok jelentették. Egy-egy szolgáltató tekintetében máig irányadó és minden fél számára előnyös szabályokat sikerült megalkotni (így például az egyszerű továbbítást a gyorsítótarolóban történő rögzítést végző szolgáltatóknál). A platform-

¹⁹ Strihó Krisztina: Elektronikus kereskedelem. *Debreceni Jogi Műhely*, 2008/1. 1–3.

²⁰ Michael Hellner: The Country of Origin Principle in the E-commerce Directive—A Conflict with Conflict of Laws? *European Review of Private Law*, Vol. 12., No. 2. (2004) 194–200.
<https://doi.org/10.54648/erpl2004015>

szolgáltatók és szolgáltatások jelentős része azonban a tárhelyszolgáltatások kategóriájába kerültek. A szolgáltatóknak ebben az esetben a jogellenes tartalomért viselt felelősség alóli mentessége megalapozásához elég volt az, hogy

„a szolgáltatóknak nincsen tényleges tudomása jogellenes tevékenységről vagy információról, és – ami a kárigényeket illeti – nincsen tudomása olyan tényekről vagy körülményekről, amelyek nyilvánvalóan jogellenes tevékenységre vagy információra utalnának; vagy a szolgáltató, amint ilyenről tudomást szerzett, haladéktalanul intézkedik az információ eltávolításáról vagy az ahhoz való hozzáférés megszüntetéséről.”²¹

Egy a szabályoknak megfelelő értesítési és eltávolítási rendszerrel dolgozó szolgáltató ez esetben elkerült minden felelősséget, tekintettel arra, hogy az irányelv egyik kulcsfontosságú szabálya az, hogy a tárhelyszolgáltatóknak nincs általános nyomkövetési kötelezettsége a felhasználók által elhelyezett tartalmak jogszerűségének ellenőrzésére. Ez a szabály jelentős védelmet nyújtott a szolgáltatók számára, de egyben szabályozási problémát is teremtett, különösen a platformokat üzemeltető szolgáltatók esetében, amelyek tevékenysége jelentősen túlmutatott az egyszerű tárhelyszolgáltatáson.²² Az Európai Unió Bírósága (EUB) több ítéletében is megerősítette, hogy az e-kereskedelmi irányelv alapján a közvetítő szolgáltatóknak, különösen a tárhelyszolgáltatóknak, nincs általános nyomkövetési kötelezettségük,²³ ahogy azt láthattuk például a SABAM v. Netlog NV (C-360/10) ügyben.

Az irányelv nem tett különbséget az egyszerű tárhelyszolgáltatók és a modern platformokat üzemeltető szolgáltatók között. Ez utóbbiak, mint például a közösségi médiaplatformok vagy online piacterek, már nem csupán tárhelyet biztosítottak, hanem aktív szerepet játszottak a tartalmak terjesztésében, moderálásában és ajánlásában. Az ilyen platformok működésével kapcsolatban egyértelműen kimutathatóvá vált a tartalommoderáció átláthatóságának hiánya, amely azért alakulhatott ki, mert az e-kereskedelmi irányelv nem tartalmazott részletes előírásokat a digitális platformok tartalommoderációs tevékenységére vonatkozóan. A platformok széles mérlegelési lehetőséget kaptak, hogy saját szabályaik alapján kezeljék a felhasználói tartalmakat. A mérlegelést részben a vonatkozó szakpolitikai fejlemények (pl. gyűlöletbeszéd elleni fellépés, kiskorúak védelme miatt) folyamatosan szigorították, miközben ezek a változtatások gyakran a véleménynyilvánítás szabadságának indokolatlan korlátozásához vezettek. Az irányelv kizárólag a jogszerűtlen tartalmakra összpontosított, és nem foglalkozott a káros, de nem feltétlenül jogellenes tartalmak (például álhírek) kezelésével. Ezzel párhuzamosan csak egy-egy speciális tartalmi szempont (pl. gyűlöletbeszéd,

²¹ E-kereskedelmi irányelv 14. cikk (1) bekezdés

²² Gergely Gosztonyi – János Bálint – Gergely Ferenc Lendvai: Public Figures and Social Media from a Freedom of Expression Viewpoint in the Recent U.S. and EU Jurisdiction. *Journalism and Media*, Vol.6, No. 1. (2025) 5. <https://doi.org/10.3390/journalmedia6010026>

²³ Gergely Gosztonyi – Ewa Galewska – Andrej Školokay: Challenges of Monitoring Obligations in the European Union’s Digital Services Act. *ELTE Law Journal*, 2024/1. 48–50. <https://doi.org/10.54148/ELTEJ.2024.1.45>

gyermekpornográfia) kapcsán követelt meg az európai szabályozás átláthatósági szempontból is releváns megfelelést a szolgáltatóktól.

Eközben a felhasználók számára nem állt rendelkezésre hatékony jogorvoslati mechanizmus a platformok moderációs és egyéb döntéseivel szemben az értesítési-eltávolítási eljárás jogintézményén kívül.²⁴ Például, ha egy felhasználói tartalmat eltávolítottak vagy felfüggesztették egy fiók működését, számos esetben tisztázatlan volt, hogy milyen módon érvényesíthetik igényeiket a felhasználók, ezzel tovább rontva a szolgáltatás nyújtásának átláthatóságát. A platformok tartalmi szűrését eközben egyre inkább uralni kezdték a háttérben működő algoritmusok, amelyekkel szemben a szabályozás még nem tudott érdemi választ adni, hiszen az irányelv kiadásának időpontjában az algoritmus alapú tartalomajánlás és a célzott hirdetések még nem voltak széles körben elterjedtek. Az irányelv ennek megfelelően nem tartalmazott előírásokat az algoritmusok működésének átláthatóságára és a felhasználók személyes adataival való visszaélések megelőzésére sem.

Ha mindez nem lett volna elég, az USA – ahol nagy arányban találhatóak a legjelentősebb platform- és keresőszolgáltatást üzemeltető vállalkozások székhelyei – szabályozásával még tovább rontotta a helyzetet. A platformszolgáltatók felelősségét leginkább a szövetségi szinten meghozott, 1996-os *Communication Decency Act* (CDA) 230. szakasza határozza meg. Ez a rendelkezés szinte teljes immunitást, felelősségmentességet biztosít a platformoknak a felhasználók által feltöltött tartalmakért, még akkor is, ha tudomásuk van a tartalom jogszerűtlenségéről.²⁵ Ugyanazon szakasz alpontja [c(2)] tartalmazza az úgynevezett „irgalmas szamaritánus” klauzulát, amely felelősségmentességet biztosít a szolgáltatóknak a polgári jogi felelősséggel szemben is, amennyiben jóhiszeműség mutatható ki a platformszolgáltatók tartalommoderálási tevékenysége kapcsán.²⁶

A jogszabályi követelménybe foglalt immunitási szabályokat még tovább erősítette az amerikai bírósági gyakorlat: a *Zeran v. America Online* ügyben (1997) a bíróság kimondta, hogy az online szolgáltatók nem felelősek a harmadik féltől származó tartalmakért, még akkor sem, ha azok károsak vagy jogellenesek. Ez az eset precedenst teremtett, amely szerint az online platformok mentesülhetnek a felelősség alól, még ha értesítették is őket a tartalom jogellenes jellegéről. A *Blumenthal v. Drudge* ügyben (1998) a bíróság rámutatott, hogy a platformok még akkor is mentesülnek a felelősség alól, ha moderálták az érintett tartalmat, amennyiben nem mutathatók ki a tartalomszolgáltatás további feltételei. A *Doe v. America Online* ügyben (2001) a Floridai Legfelső Bíróság, valamint a *Doe v. MySpace* ügyben (2008) a Fifth Circuit (az Egyesült Államok ötödik körzeti fellebbviteli bírósága) megerősítette, hogy a szolgáltató

²⁴ Gyetván Dorina: A Digital Services Act és a Facebook Oversight Board szerepe a jogorvoslat biztosításában a közösségi média vonatkozásában. *In Medias Res*, 2023/2. 192–195.
<https://doi.org/10.59851/imr.12.2.9>

²⁵ Michelle Jee: New Technology Merits New Interpretation: An Analysis of the Breadth of CDA Section 230 Immunity. *Houston Business & Tax Law Journal*, Vol. 13., No. 2. (2013) 186–193.

²⁶ Gosztonyi Gergely: Az internetes tartalomszabályozással kapcsolatos új gondolkodási irányok az Amerikai Egyesült Államokban. *Miskolci Jogi Szemle*, 2021/4. 40–54.
<https://doi.org/10.32980/MJSz.2021.4.1257>

nem felelős a harmadik féltől származó jogellenes tartalmakért, mivel a 230. szakasz védelmet nyújt az ilyen helyzetekben. A bírósági joggyakorlat odáig is eljutott, hogy 2014-ben a *Jones v. Dirty World Entertainment Recordings LLC* ügyben a bíróság meghatározta, hogy a platform üzemeltetése nem minősül tartalomszerkesztési tevékenységnek, és ezért a 230. szakasz védelmet nyújtott számára.

Az elmúlt években az USA-ban számos reformjavaslat látott napvilágot, amelyek a CDA 230-as szakasz átfogó felülvizsgálatát szorgalmazták,²⁷ azonban ezek sem tudták általános jelleggel áttörni a felelősségmentességet, csak egy-egy speciális területen.²⁸ Az amerikai felelősségmentességet megállapító szabályok és joggyakorlat miatt pedig csak az európai szabályozás fektethette le azokat a rendelkezéseket, amelyek az MI által végzett tartalommoderációs tevékenység alapvető követelményeinek tekinthetünk.

Az elektronikus kereskedelmi irányelv kiadása és a digitális szolgáltatásokról szóló rendelet elfogadása közötti időszakban éppen ezért több olyan próbálkozás is történt európai szinten, amely a platformszolgáltatók működésének szabályozására irányult, különösen az átláthatóság szempontjából. Ezek ugyan csak részeredményeket jelentettek, de a DSA megszületésének előképét jelentették.

E folyamat egyik meghatározó állomása volt a német *Netzwerkdurchsetzungsgesetz* (NetzDG) elfogadása 2017-ben,²⁹ amely a nagy platformszolgáltatók számára előírta a rendszeres jelentéstételi kötelezettséget a jogszerűtlen tartalmak kezelésével kapcsolatban (2. cikk), és eljárási szabályokat határoltak le e tartalmakkal kapcsolatos panaszok kezelésére (3. cikk). A NetzDG a DSA-ban megjelenő transzparencia-kötelezettségek egyik mintájául szolgált, különösen a bejelentések, eltávolítási arányok és moderálási döntésekről való jelentéstétel témaköreiben. Franciaországban hasonló céllal született meg a *Loi Avia*,³⁰ ami azonban a francia alkotmánybíróság döntése értelmében végül nem lépett hatályba. A javaslat célja a terrorizmushoz, gyűlöletbeszédhez kapcsolódó és erőszakot bemutató tartalmak gyors eltávolítása volt, és kötelező eltávolítási határidőket írt volna elő. A jogszabály az alkalmazás hiányának ellenére is jelentős diskurzust generált a tartalommoderációs kötelezettségek és az átláthatóság kérdéseiről.³¹

Ezekkel párhuzamosan az Európai Bizottság is felismerte a szabályozási szükséglet az átláthatósági kérdések tekintetében felmerülő kihívások kapcsán több alkalommal

²⁷ Michael L. Rustad – Thomas H. Koenig: The Case for a CDA Section 230 Notice-and-Takedown Duty. *Nevada Law Journal*, Vol. 23., No. 2. (2023) 533.

²⁸ Ld. pl. a szexuális jellegű jogellenes tartalmak terén hozott változást az úgynevezett *FOSTA-SESTA* (2018) jogszabálysomag, amely kivételt teremtett a szakasz alkalmazása alól az online szexkereskedelemmel kapcsolatos tartalmak esetében.

²⁹ *Netzwerkdurchsetzungsgesetz vom 1. September 2017* (BGBl. I S. 3352), das zuletzt durch Artikel 29 des Gesetzes vom 6. Mai 2024 (BGBl. 2024 I Nr. 149) geändert worden ist.

³⁰ *Assemblée Nationale, Proposition de loi n° 388, adoptée par l'Assemblée nationale, en nouvelle lecture, visant à lutter contre les contenus haineux sur internet*, 13 Mai 2020

³¹ Dominik Steiger: Protecting Democratic Elections Against Online Influence via “Fake News” and Hate Speech—The French *Loi Avia* and *Loi No. 2018–1202*, the German Network Enforcement Act and the EU’s Digital Services Act in the Light of the Right to Freedom of Expression. In: Stephanie Schiedermaier – Alexander Schwarz – Dominik Steiger (szerk.): *Theory and practice of the European Convention on Human Rights*. Baden-Baden, Nomos Verlagsgesellschaft, 2021. 175.
<https://doi.org/10.5771/9783748923503-165>

is,³² sürgette a szolgáltatókat, hogy önkéntes alapon hozzanak létre átláthatóságra épülő moderációs eljárásokat. Ezek gyakorlati hatása azonban – kötelező előírások hiányában – korlátozott maradt.³³

3. A DSA vonatkozó követelményei

A DSA egyik legfontosabb szabályozási újítása, hogy az e-kereskedelmi irányelv széles felelősségmentességi szabályai helyett a felelősségmentességet leszűkítő és a kockázatarányos szabályozás keretei között,³⁴ átláthatósági mechanizmusokat ír elő a tartalommoderáció minden szakaszára. Ezen irányultság tükröződik a jogszabályalkotói célokat felvázoló és szabályozási megoldásokat körvonalázó preambulumbekendésekben is, amelyek körvonalazzák a rendelet normatív szabályozását is.

A DSA preambuluma nem pusztán deklarálja a jogalkotói célokat, a szabályozás előzményeit és azokat a megfontolásokat, amelyek a szabályozási logika megalkotásához vezettek: a benne foglalt szabályozási elvek kötöttséget is hordoznak a rendelet értelmezése során és ez igaz a transzparenciára vonatkozó követelmények tekintetében is. Az európai uniós szabályozás fejlődéstörténete alapján is megállapítható, hogy az egyes preambulumbekendések gyakorlati jelentőséggel bírnak a jogalkalmazói értelmezés során és az Európai Unió Bíróságának joggyakorlatában. A transzparencia e jogalkotói környezetben nem csupán általános elvárásként, hanem a szolgáltatók működési kötelezettségeit meghatározó, horizontális szabályozási elemként jelenik meg. Olyan jelentős szabályozási elemként azonosítható, amely áthatja az egész szabályozás követelményrendszerét, mert a DSA majd minden részletszabálya tartalmaz olyan elemeket, amelyek a platformszolgáltatások átláthatóságával összefüggésben vannak. Ezt arra tekintettel lehet kijelentni, hogy a DSA szabályozását az európai jogalkotó a rendelet preambuluma alapján olyan kiszámítható, biztonságos és megbízható online környezet megteremtése érdekében alkotta meg,³⁵ amely képes kezelni a jogellenes tartalmak gyors terjedéséből eredő társadalmi kockázatokat, ugyanakkor biztosítja a szolgáltatások igénybe vevőinek alapvető jogait is. Ez immanensen magába foglalja azt

³² Ld. C/2018/1177 A Bizottság (EU) 2018/334 ajánlása (2018. március 1.) az illegális online tartalom hatékony kezelésére irányuló intézkedésekről. HL L 63., 2018.3.6., 50–61. o.

³³ Evangelia Psychogiopoulou: The Fight Against Digital Hate Speech: Disentangling the EU's Regulatory Approach and Hurdles. *German Law Journal*, Vol. 25., No. 7. (2024) 1190. <https://doi.org/10.1017/glj.2024.26>

³⁴ Sara Tommasi: Risk-Based Approach in the Digital Services Act and in the Artificial Intelligence Act. In: Sara Tommasi (szerk.): *The Risk of Discrimination in the Digital Market: From the Digital Services Act to the Future*. Cham, Springer Nature Switzerland, 2023. 73–76. https://doi.org/10.1007/978-3-031-43640-6_5

³⁵ DSA (9) preambulumbekendés. Összhangban: COM/2020/67 final A Bizottság Közleménye az Európai Parlamentnek, a Tanácsnak, az Európai Gazdasági és Szociális Bizottságnak és a Régiók Bizottságának – Európa digitális jövőjének megtervezése. 2020.02.19.; Szöke Gergely László: A közösségi oldalak szabályozási problémái. In: Kis Kelemen Bence – Mohay Ágoston (szerk.): *A technológiai fejlődés jogi kihívásai. Kézikönyv a jogalkotás és jogalkalmazás számára*. Pécs, PTE ÁJK, 2021. 105–107.

is, hogy a szabályozott szolgáltatásokat olyan keretek közé szorítja,³⁶ amelyek átláthatóbbá teszik a szolgáltatás igénybevételi feltételeit, a működés egyes kritikus kérdéseit. A rendelet hatálya az összes közvetítő szolgáltatásra, így a tárhelyszolgáltatókra, azon belül az online platformokra és online piacterekre is kiterjed, de kiemelten kezeli az online óriásplatformokat (VLOP) és a nagyon népszerű keresőprogramokat (VLOSE), amelyeknek különösen nagy hatásuk van az alapvető jogokra. A szabályozás ugyan fenntartja az e-kereskedelmi irányelvben szereplő, általános monitorozási vagy aktív tényfeltárási kötelezettségek hiányát megállapító rendelkezést (8. cikk), de ezt áttöri a rétegzett szabályozáson keresztül a legnagyobb és legjelentősebb hatású szolgáltatókra vonatkozó kötelezettségek körében.

A platformszolgáltatások vonatkozásában éppen e rétegzett szabályozásra tekintettel szét kell választani azokat a rendszerint *ex ante* jellegű átláthatósági kötelezettségeket, amelyek minden közvetítő szolgáltatás irányába alkalmazandóak lesznek és azokat, amelyek csak az online óriásplatformokra és a nagyon népszerű keresőprogramokra lesznek irányadóak. A transzparenciához kapcsolódó követelmények több, egymással párhuzamosan érvényesülő szabályozási csomópont köré épülnek, amelyek az átláthatóság követelményrendszerének mintegy pilléreiként épülnek bele a szabályozásba.

3.1. Szerződési feltételek átláthatósági vonatkozásai

Egyik legfontosabb új szabályozási elem az átláthatóság szempontjából a platformszolgáltatók felhasználási feltételeire vonatkozik, amely kifejezetten szigorú közzétételt és transzparenciát ír elő, speciális tájékoztatási kötelezettséggel egybekötve.³⁷ E feltételek azért bírnak jelentőséggel, mert a felhasználók legfontosabb döntése, amikor a szolgáltatás igénybevételéről döntenek, s ebben a pozícióban különös jelentősége van átláthatósági szempontból is, hogy milyen információk állnak rendelkezésükre és mennyire közérthető az erre vonatkozó tájékoztatás.³⁸ A következetes tájékoztatással biztosítható ugyanis, hogy a felhasználók előre tudják, milyen tartalmak megosztása engedélyezett vagy tiltott, és a moderáció kiszámíthatóbb, jogszerűbb keretek között történjen. A jogalkotó szándéka továbbá a túlzottan homályos vagy önkényes platformszabályok visszaszorítása,³⁹ mivel ezek korábban a szólásszabadság indokolatlan korlátozásához

³⁶ Önmagában azzal is, hogy szabályozott kereteket hoz létre olyan területen, ahol korábban nem voltak szabályok és így nem voltak átláthatóak a felhasználók számára a rá vonatkozó követelmények. Ld. Ellen Wauters – Verónica Donoso – Eva Lievens: *Optimizing transparency for users in social networking sites. Info/Digital Policy, Regulation and Governance*, Vol. 16., No. 6. (2014) 18. <https://doi.org/10.1108/info-06-2014-0026>

³⁷ Gellén Klára: A DSA fogyasztókat érintő egyes rendelkezései a fogyasztói bizalom és az átláthatóság erősítése érdekében. In: Koltay András – Szikora Tamás – Lapsánszky András (szerk.): *A vadnyugat vége? Tanulmányok az Európai Unió platformszabályozásáról*. Budapest, ORAC, 2024. 114. https://doi.org/10.59851/9789632586328_06

³⁸ Giuseppe A. Veltri et al.: The impact of online platform transparency of information on consumers' choices. *Behavioural Public Policy*, Vol. 7. No. 1. (2023) 75–79. <https://doi.org/10.1017/bpp.2020.11>

³⁹ Zódi Zsolt: Átláthatósági és indokolási követelmények az európai platformjogban. *In Medias Res*, 2023/2. 14–17. <https://doi.org/10.59851/imr.12.2.1>

vezettek (pl. „árnyéktiltásokhoz”, amikor a felhasználó tudta nélkül korlátozták a tartalma láthatóságát).⁴⁰ Már a preambulumbekendések [(45)–(48)] lefektetik annak kerekeit, hogy a szolgáltatóknak – különösen, ha automatizált eszközöket, algoritmusokat vagy mesterséges intelligencia alapú megoldásokat is alkalmaznak – a felhasználási feltételekben egyértelműen fel kell tüntetniük, hogy milyen szempontok alapján korlátozhatják a szolgáltatás igénybevételét vagy távolíthatnak el tartalmat, s ez egyszerre bír jelentőséggel a tartalommoderáció tekintetében és minden olyan irányelv, eljárás, technikai megoldás kapcsán is, amelyet ennek során alkalmaznak. E körben kell értékelniük és végül a szerződési feltételek között szerepeltetniük a szolgáltatóknak, ha MI alapú technológiai megoldásokat alkalmaznak a moderációs tevékenység során, nyilvánvalóan azért, mert a felhasználók így érzékelhetik e technológiai megoldások használatából eredő nagyobb vagy eltérő jellegű kockázatokat, amikor döntést hoznak a szolgáltatás igénybevételéről. Valójában ugyanez lesz az alapja az átláthatósági jelentési kötelezettségek (24. cikk) és hirdetési- és ajánlórendszerek átláthatósági követelményeinek (26–27. cikk) is.

Az MI felhasználása tekintetében ezen a ponton kapcsolódik össze a később kiadásra került AI Act és a DSA követelményrendszere, mert a szerződési feltételekben mind a tartalommoderáció módszertanát, mind pedig az alkalmazásra kerülő MI rendszerek úgynevezett műszaki leírását szerepeltetni kell, annak érdekében, hogy a szolgáltató egyszerre tudja teljesíteni a két szabályozás követelményeit. A szerződési követelmények tisztázása e tekintetben mutat kógens tartalmi elemeket is,⁴¹ hiszen a felhasználási feltételeik tartalmára, alkalmazására és érvényesítésére vonatkozóan úgy kell megállapítani a szabályokat, hogy azok a transzparenciát, a felhasználóvédelmet és a tisztességtelen és önkényes megoldások elkerülését szolgálják. Azt már helyütt érdemes közreadni, hogy az átláthatóság növelése nem mindig a legegyszerűbben teljesíthető ezen keretek között sem – a szerződési feltételek egyértelmű, érthető és előzetes közlése mindenképpen átláthatóságot növelő hatással bír, azonban az MI rendszerek összetett működése nehezítheti az alkalmazott technikák közérthető és mindenre kiterjedő leírását. E követelmény teljesítése még tovább nehezedhet, ha figyelembe vesszük azt is, hogy a sérülékeny csoportok tekintetében, illetve kiskorúakat célzó vagy nagyrészt általuk használt szolgáltatások esetén a feltételek magyarázatát könnyen érthetővé kell tenni a számukra.⁴² A jogi nyelvezetű feltételeket egyszerre részletesen és mégis közérthetően kell megfogalmazni, úgy, hogy minden felhasználó – akár kiskorú vagy jogban járatlan – megértse. Különösen olyan technológiai megoldásoknál, mint a mesterséges intelligencia alkalmazások, egyáltalán nem egyszerűen megvalósítható a rendkívüli összetett döntési mechanizmus közérthető és egyszerű megfogalmazása.

⁴⁰ David Nosák: The DSA Introduces Important Transparency Obligations for Digital Services, but Key Questions Remain. Center for Democracy & Technology, June 18, 2021. <https://tinyurl.com/2f564bvh>; Jean De Meyere: The Digital Services Act: Towards more transparency for content moderation. KU Leuven CITIP, 13 February, 2024. <https://tinyurl.com/2986y7aa>

⁴¹ Ld. DSA (45) preambulumbekendés.

⁴² João Pedro Quintais – Naomi Appelman – Ronan Ó. Fathaigh: Using terms and conditions to apply fundamental rights to content moderation. *German Law Journal*, Vol. 24. No. 5. (2023) 893–894. <https://doi.org/10.1017/glj.2023.53>

Nem egyértelmű például, mit tekint majd a Bizottság „világos és egyértelmű” nyelvnek,⁴³ és ez szolgáltatótípusonként is eltérhet, ami az első DSA 66. cikke szerinti eljárások esetében jogalkalmazási nehézségeket fog okozni.

3.2. Bejelentési és cselekvési mechanizmusok átláthatósági vonatkozásai

A DSA 16. cikkében szabályozza a bejelentési és intézkedési mechanizmusokat. Ezek arra szolgálnak, hogy minden hosting jellegű szolgáltató (tárhelyszolgáltató vagy online platform) hatékony és felhasználóbarát eljárást biztosítson a jogellenes tartalmak bejelentésére és eltávolítására. Ez a rendelkezés egységes keretet teremt az EU-ban a felhasználói bejelentések (*notice*) kezelésére és az arra adott platformintézkedésekre (*action*).⁴⁴ A jogalkotó szándéka a jogellenes tartalmak gyorsabb eltávolítása a bejelentések nyomán, miközben biztosítja, hogy a bejelentők visszajelzést kapjanak a panaszukról. A cikk összhangban van a korábbi „értesítés és eltávolítás” elvével, de azt explicitté és kötelező erejűvé teszi minden szolgáltató számára. Bejelentés esetén a szolgáltató dokumentáltan „tudomást szerez” a jogsértésről, így köteles lépni (ellenkező esetben elveszítheti a felelősség alóli mentességét).⁴⁵ Az átláthatóságot növeli a cikk előírásai szerint, hogy a tárhelyszolgáltatóknak (beleértve az online platformokat) könnyen hozzáférhető és felhasználóbarát elektronikus bejelentőfelületet kell működtetniük és ezen keresztül minden bejelentést időben, kellő gondossággal, az önkényesség kizárásával és tárgyilagos módon kell feldolgozniuk. Ha automatizált eszközt használ a bejelentések feldolgozására vagy a döntésre, azt jeleznie kell a bejelentőnek küldött értesítésben – itt ismét visszatérhet az MI szerepe az átláthatóságot biztosító folyamatokban.

A bejelentési mechanizmusokon és belső panaszkezelő rendszeren keresztüli bejelentések tekintetében a DSA 17. cikkében rendezi a döntések indoklási kötelezettségére vonatkozó szabályait. E cikk a platformok döntéseinek átláthatóságát és elszámoltathatóságát hivatott erősíteni azzal, hogy minden lényeges moderációs intézkedésről indokolást kell adniuk az érintett felhasználónak. Ez a rendelkezés a korábbi gyakorlat hiányosságaira reagál: sokszor előfordult, hogy a felhasználók nem kaptak semmilyen magyarázatot tartalmaik törlése vagy fiókjaik letiltása esetén, ami kiszolgáltatottságot és a szólásszabadság csorbulását eredményezte.⁴⁶ A részletes indoklás követelménye arra ösztönzi a platformokat, hogy körültekintően és ne önkényesen járjanak el, hiszen

⁴³ Nataliia Filatova-Bilous – Tetiana Tsuvina – Bohdan Karnaukh: Digital Platforms’ Practices on Content Moderation: Substantive and Procedural Issues Proposed by DSA. In: Mykola Nechyporuk – Volodymir Pavlikov – Dmytro Krytskyi (szerk.): *Integrated Computer Technologies in Mechanical Engineering – 2023*. Cham, Springer Nature Switzerland, 2023. 198–200. https://doi.org/10.1007/978-3-031-60549-9_16

⁴⁴ João Pedro Quintais et al.: *From the DMCA to the DSA: A Transatlantic Dialogue on Online Platform Regulation and Copyright*. SSRN, 17 Jun 2024. 34–37. <https://dx.doi.org/10.2139/ssrn.4860225>

⁴⁵ Daniel Holznapel: How to Apply the Notice and Action Requirements under Art. 16 (6) sentence 1 DSA—Which action actually? *Computer Law Review International*, Vol. 25., No. 6. (2024) 172–173.

⁴⁶ Paddy Leerssen: An end to shadow banning? Transparency rights in the Digital Services Act between content moderation and curation. *Computer Law & Security Review*, Vol. 48., No. 105790. (2023) 6–8. <https://doi.org/10.1016/j.clsr.2023.105790>

döntéseiket meg kell tudniuk indokolni. A cikk előírja: a felhasználók számára lehetővé kell tenni, hogy megértsék a döntés okát és éljenek a jogorvoslati lehetőségeikkel, ha vitatják azt. A 17. cikk (4) bekezdése előírja, hogy az indokolásnak amennyire csak lehet, pontosnak és specifikusnak kell lennie, ugyanakkor közérthető nyelvezetűnek, ami újbóli nehézség elé állíthatja a platformszolgáltatókat, de ennek segítségével biztosítható érdemben a felhasználók átláthatóságérzetének valódi kialakulása. Felmerülhet azonban a fent már említett, az MI jellegzetességéből adódó probléma: nehezen megoldható lehet az MI-vel támogatott moderációs tevékenység utólagos indoklása és annak közérthető közreadása.⁴⁷

A jogorvoslatok körét kívánja bővíteni a jogalkotó a DSA 20. cikkével, amely a belső panaszkezelési rendszerrel, és 21. cikkel, amely a bíróságon kívüli vitarendezés részleteivel foglalkozik. Átláthatósági szempontból ezeknek azért van szerepe, mert egyrészt fel tudják oldani a korábbiakban joghatósági problémák miatt jelentkező igényérvényesítési zavarokat, másrészt közelebb tudják hozni a tartalommoderációs döntéseket és az arra vonatkozó jogorvoslatokat az érintettekhez. Azzal, hogy a belső panaszkezelés következetesen dokumentálásra kerül, nyomkövethető a folyamat, a felhasználó számára egyértelmű visszajelzések küldése önmagában átláthatóság növelő hatását, az ezt kiegészítő, független fórum által lefolytatott eljárás pedig tovább növelheti a felhasználók bizalmát és szubjektív átláthatóságérzetét a tartalommoderációval, a döntésekkel és önmagában a jogorvoslati folyamattal szemben is.

3.3. Átláthatósági jelentések

Átláthatósági szempontból látszólag az egyik legnagyobb jelentőségű követelményt hozza létre a DSA 15. cikke, meghatározva az átláthatósági jelentéstételi kötelezettséget, amely elsődlegesen a tartalommoderációs tevékenységről történő beszámolást helyette javaslat: beszámoló elkészítését jelenti. A cikk azt a célt szolgálja, hogy a szolgáltatók rendszeres időközönként számot adjanak tartalommoderációs tevékenységükről, ezzel növelve a platformok elszámoltathatóságát és a közbizalmat.⁴⁸ Az éves (vagy a nagyobb szolgáltatóknál gyakoribb) átláthatósági jelentések révén a felhasználók, a hatóságok és a civil szféra átfogó képet kaphat arról, hogyan alkalmazzák a platformok a moderációs szabályokat, milyen mértékben távolítanak el illegális vagy szabálysértő tartalmakat, és hogyan működnek az új panasz eljárások. E jelentésben kell bemutatnia a szolgáltatóknak azt, hogy az általa alkalmazott – MI alapú és más jellegű – technológiai megoldások mennyire hatékonyak, milyen problémákkal szembesült a vizsgált időszakban. A jogalkotó szándéka az, hogy az eddig önkéntes alapon közzétett átláthatósági jelentéseket kötelezővé és összehasonlíthatóvá tegye, egységes kategóriák mentén. Így a jelentések elősegítik a tartalommoderációs gyakorlatok ellen-

⁴⁷ Yifat Nahmias – Maayan Perel: The oversight of content moderation by AI: impact assessments and their limitations. *Harvard Journal on Legislation*, Vol. 58., No. 1. (2021) 145.

⁴⁸ Marta Maroni: “Mediated transparency”: The Digital Services Act and the legitimisation of platform power. In: Maarten Hillebrandt – Päivi Leino-Sandberg – Ida Koivisto (Szerk): *(In)Visible European Government*. London, Routledge, 2023. 312–326.

őrzését és a problémás területek azonosítását (pl. ha egy platformhoz tömegesen érkeznek bizonyos típusú panaszok).⁴⁹ Emellett a 15. cikk fontos szerepet játszik a hatósági felügyelet megalapozásában is: a szolgáltatók nyilvános adatszolgáltatása megkönnyíti a szabályok betartásának monitorozását. A 15. cikk (1) bekezdése részletezi, hogy a jelentésnek milyen adatokat kell tartalmaznia. Hozzá kell azonban tenni, hogy az átláthatóság szempontjából lényeges szereplők, a felhasználók nézőpontjából az átláthatósági jelentés nem közvetlenül szubjektív átláthatóságot növelő kötelezettség, mert a felhasználóknak csak elenyésző része olvassa el ezeket, sokkal gyakrabban foglalkozik a platform általánosan használt felületeivel és az azokon keresztül biztosított szolgáltatásokkal. A DIGITALEUROPE által készített 2024. januári jelentés szerint a DSA által előírt átláthatósági jelentések sablonjai elsősorban a kutatók és hatóságok számára készülnek, nem felhasználóbarátok, sem nem könnyen érthetőek az egyéni szolgáltatások ügyfelei vagy felhasználói számára. A jelentés kiemeli, hogy még a kvalitatív jelentési sablon sem alkalmas a legtöbb felhasználó számára, mivel olyan részleteket tartalmaz, mint a jogsértő tartalmak detektálási módszerei, a felhasználási feltételek felülvizsgálatai, a moderátorok képzése vagy az automatizált jelentések különböző adatai, amelyek nem nyújtanak tömör és érthető összefoglalókat a felhasználók számára.⁵⁰

A DSA réteges szabályozása alapján az online platformokra további átláthatósági jelentési követelmények kerülnek meghatározásra. Célja, hogy a 15. cikk szerinti általános tartalommoderációs jelentést kiegészítse olyan információkkal, amelyek az új DSA mechanizmusok hatékonyságát mutatják, illetve segítsék az online óriásplatformok (VLOP-ok) azonosítását.⁵¹ A jelentések e követelmények alapján visszajelzést adnak abban is, hogyan működnek a bevezetett jogorvoslati és visszaélés-elleni eszközök (pl. mennyi peren kívüli vitarendezés van, hány felhasználót függesztettek fel visszaélésért). Ezen túlmenően előírja a felhasználói számok közzétételét, ami a VLOP minősítés egyik alapja (33. cikk szerinti küszöb). Végül a 24. cikk (5) bekezdése egy fontos átláthatósági újítást vezet be: a platformoknak a tartalommoderációs indoklásokat be kell küldeniük egy uniós nyilvántartásba, hogy az kutatható legyen. A szabályozási cél itt a külső ellenőrzés és kutatás elősegítése – a Bizottság által kezelt adatbázis révén a civil társadalom, tudósok, újságírók is betekintést nyerhetnek a platformok konkrét moderációs döntéseibe.⁵² Egyúttal azonban nehezített pályát jelent az MI alkalmazása során: a mesterséges intelligencia alkalmazásokat rendszerint döntésre vagy döntéselőkészítésre alkotják meg, kevésbé ezek indoklására, amely immánensen nehezíti az indoklások kiadását.

⁴⁹ Aleksandra Urman – Mykola Makhortykh: How transparent are transparency reports? Comparative analysis of transparency reporting across online platforms. *Telecommunications policy*, Vol 47., No. 3. 102477. (2023) 11–12. <https://doi.org/10.1016/j.telpol.2022.102477>

⁵⁰ Digital Services Act transparency reports: focus on simplicity and usefulness. *DigitalEurope*, 22 January 2024. <https://tinyurl.com/mr32tnr4>

⁵¹ Kasia Söderlund: Regulating high-reach AI: On transparency directions in the Digital Services Act.” *Internet Policy Review*, Vol. 13. No. 1. (2024) 3. <https://doi.org/10.14763/2024.1.1746>

⁵² Gautam Kishore Shahi et al.: *A Year of the DSA Transparency Database: What it (Does Not) Reveal About Platform Moderation During the 2024 European Parliament Election*. <https://doi.org/10.48550/arXiv.2504.06976>

A Bizottság a követelmények teljesítésének elősegítésére 2024 novemberében kiadott egy végrehajtási rendeletet,⁵³ amely megalapozza az átláthatósági jelentések egységes kialakítását, módszertanát, és annak érdekében, hogy e jelentések valóban széles körben elérhetővé, tudományosan elemezhetővé (akár számaadataiban is) váljanak, e végrehajtási rendelet mellékletében közzétette az átláthatósági jelentések elvárt formáját. Ebben azonban nem említi külön a Bizottság, hogy értelmezése alapján további kötelezettségek terhelnék az MI alapú technikát alkalmazó szolgáltatókat, a jelentés tartalmának követelményeit e tekintetben a DSA adja. Az érintett szolgáltatóknak e jelentéseiket 2025 júliusáig kell eljuttatni a Bizottság részére és közzétenni.⁵⁴

3.4. Hirdetési- és ajánlórendszerek átláthatósága

A DSA 26. cikke eközben az online hirdetések átláthatóságával foglalkozik, mely közvetve a tartalommoderációhoz is kapcsolódik (hiszen a reklámok is tartalomként jelennek meg, és sok panasz épp a nem jelölt, megtévesztő hirdetések miatt merül fel). A téma azért is fontos számunkra, mert napjainkra a hirdetési tevékenység központi alapját is az MI által támogatott folyamatok adják, megkerülhetetlen a működésük a modern hirdetési rendszerekben.⁵⁵ A szabályozó célja, hogy a felhasználók egyértelműen felismerjék a hirdetéseket, tudják, ki áll mögöttük és miért látják őket. Ezzel küzd a rejtett reklám és manipuláció ellen, és javítja a szolgáltatás átláthatóságát. A platformok eddig is alkalmaztak jelöléseket („Szponzorált” címke stb.), de a DSA ezt kötelező normává emeli, és kiegészíti a célzási paraméterekre vonatkozó információkkal. Az is cél, hogy a felhasználók visszajelzést adhassanak (pl. jelölhessék a saját tartalmukat reklámként), így a platform ne csak maga próbálja felismerni. Végül a cikk tilt egy bizonyos gyakorlatot: érzékeny adatokon alapuló célzott hirdetést nem lehet nyújtani, ami a felhasználók magánszféráját védi. Ezek az adatok jellemzően a platform által felmérhető emberi viselkedésből levont következtetések, amelyek jelentős mennyiségű különleges személyes adatot is tartalmazhatnak, s kezelésük az adatvédelmi követelmények alapján jogszerűtlen a célok és jogalapok precíz meghatározásának hiányába. Tömeges és MI rendszerekkel való kezelésük pedig tovább ronthatja a szolgáltatások átláthatóságát és a felhasználók szubjektív átláthatóságérzetét.⁵⁶

A DSA 27. cikke a fenti szabályokat kiegészítve a platformok tartalomajánló algoritmusainak (*recommender systems*) működésébe enged bepillantást a felhasználók

⁵³ 2024/2835 végrehajtási rendelet a közvetítő szolgáltatók és az online platformot üzemeltető szolgáltatók (EU) 2022/2065 európai parlamenti és tanácsi rendelet szerinti átláthatósági jelentési kötelezettségeire vonatkozó mintadokumentumok meghatározásáról. HL L, 2024/2835, 2024.11.05. 1-48. o.

⁵⁴ European Commission: Implementing regulation laying down templates concerning transparency reporting obligations. <https://tinyurl.com/266zb2vb>

⁵⁵ Elena Izyumenko – Martin Senftleben: Online behavioural advertising, consumer empowerment and fair competition: are the DSA transparency obligations the right answer?. *Journal of European Consumer and Market Law*, Vol. 14., No. 2. (2025) 46. <https://doi.org/10.2139/ssrn.4729118>

⁵⁶ Bram Duivenvoorde – Catalina Goanta: The regulation of digital advertising under the DSA: A critical assessment. *Computer Law & Security Review*, Vol. 51. No. 105870. (2023) 8–10. <https://doi.org/10.1016/j.clsr.2023.105870>

számára, és bizonyos fokú felhasználói kontrollt biztosít felettük. A cél az, hogy a felhasználók tudatában legyenek, mi alapján rangsorolja vagy javasolja nekik a platform a tartalmakat (pl. hírfolyam, videóajánló), és lehetőségük legyen ezen preferenciáikat alakítani. Ez tulajdonképpen az algoritmusok „fekete doboz”-problémája. Korábban a felhasználók semmit, vagy csak nagyon keveset tudtak arról, miért pont azt a tartalmat látják, amit, s ez aggályokat vet fel a tájékozódáshoz való jog és a véleménypluralizmus szempontjából. Az alapvető problémát ez a követelmény nem tudja feloldani,⁵⁷ de legalább láthatóvá teszi, hogy milyen feladatokra, milyen adatok alapján alkalmazza a platform az ajánlórendszert. A szabályozó kimutatható szándéka, hogy a platformok a főbb rendezési szempontokat nyilvánossá tegyék, amelynek bár megkérdőjelezhető mértékű a bizalomépítő hatása, de legalább az elszámoltathatóságot segíti (ha a paramétereiből kiderülne például, hogy előnyben részesít bizonyos politikai nézeteket, azt a közösség számonkérheti). Továbbá a felhasználók kezébe ad némi döntési lehetőséget: ha van más ajánlási opció (pl. időrendi sorrend), választhassanak és részben kimaradhatnak az ajánlórendszerbe foglalt szabályok egy része alól.⁵⁸ Mindez önmagában is növelheti a felhasználók átláthatóságérzetét. Az elszámoltathatóságot pedig növeli, hogy a felhasználókon túl az Európai Bizottság és tagállami hatóságok, így a digitális szolgáltatási koordinátor felügyeleti lehetősége is körvonalazásra kerül: a kockázatok és lehetséges károk felméréséhez szükséges adatokhoz, az algoritmikus tartalommoderálási rendszerek, az ajánlórendszerek vagy a hirdetési rendszerek pontosságára, működésére és tesztelésére, illetve a tartalommoderálás vagy az e rendelet értelmében vett belső panaszkezelési rendszerek folyamataira és kimeneteire vonatkozó adatokhoz a hatóságok hozzáférhetnek, segítve a platformok átlátható működése feletti kontrollt.

4. Következtetések és javaslatok

A tanulmányba foglalt elemzést követően több következtetést és javaslatot szeretnék megfogalmazni a DSA tartalommoderációs szabályai és a mesterséges intelligencia felhasználásának kapcsolatáról, figyelemmel a megfogalmazott kutatási kérdésre.

A DSA érdemi változást hozott az online platformok tartalommoderációjának átláthatóságára vonatkozó szabályozásban. Az eddig csak társadalmi elvárások kereszt-tüzében, önkéntesen végzett tevékenységek e szabályozás segítségével általánosan alkalmazandó kereteket kaptak, amely nagy előrelépést jelent. A részletes indokolási kötelezettség, a DSA tartalommoderációs átláthatósági adatbázisa, a rendszeres átláthatósági jelentések előírása és tartalmuk szabályozása mind-mind e rendszer fontos elemeit jelentik. Mindez nemcsak a hatósági felügyelet biztosítása szempontjából fontos: a jelentések önmagukban is visszajelzést adnak a felhasználóknak és maguknak a

⁵⁷ Miatán az algoritmusok működéséből az adódik, hogy a különböző szempontok figyelembe vétele és általánosságban a döntéshozatali folyamatok nem megismerhetők, hanem azok az algoritmus akár valós időben változó belső döntéshozatali logikája szerint alakulnak.

⁵⁸ Lendvai Gergely Ferenc: Taming the Titans? – Digital Constitutionalism and the Digital Services Act. *ESSACHESS – Journal for Communication Studies*, Vol. 17., No. 2. (2024) 175–176. <https://doi.org/10.21409/0M3P-A614>

platformokat üzemeltető szolgáltatásoknak automatikus moderációs eszközeik pontosságáról, hibaarányáról, a humán moderátorok munkájának arányáról.

Az is megállapítható, hogy a vezető online óriásplatformok már a DSA hatálybalépését megelőzően komolyan vették a szabályozást és elkezdtek átalakítani rendszereiket, hogy megfeleljenek az új szabályoknak.⁵⁹ A legtöbb platform, még mielőtt a DSA követelményei szerint kötelező lett volna, frissítette a felhasználási feltételeit és jelentősen bővítették a moderációs adatok közzétételét. 2023 őszétől a kijelölt nagy platformok sorra publikálták első átláthatósági jelentéseiket, szinte kivétel nélkül.⁶⁰ A legnagyobb szolgáltatók közül a Google egy átfogó Átláthatósági Központot hozott létre, ahol szolgáltatásonként ismerteti moderációs irányelveit és gyakorlatait.⁶¹ A TikTok jelentős többleterőforrásokat fordít moderációra: 2024 első felében a szabálysértő tartalmak 80%-át már automatikus eszközökkel távolították el (szemben az előző évi 62%-kal), miközben továbbra is több mint 6000 főt foglalkoztatnak EU-nyelvi moderátorként minden uniós tagállam nyelvét lefedve. Ezzel jól bemutatják a szolgáltatók központi stratégiáját: az MI alapú tartalommoderációt a kontextus értékelése során kiegészítik emberi felügyelettel, ezzel növelve a moderációs tevékenység hatékonyságát.

Jelenleg semmilyen jel nem utal arra, hogy a DSA miatt a platformoknak csökkenteniük vagy leépíteniük kellene a mesterséges intelligencia alkalmazását a moderációban. A DSA nem tiltja az automatizált moderációt – épp ellenkezőleg, számol annak használatával, de átláthatóvá és látható felelőségek mellett kívánja szabályozni azt. A rendelet előírja az automatizált eszközök teljesítményének mérését és közlését (pontosság, hibaarány),⁶² ami arra ösztönzi a platformokat, hogy tovább javítsák algoritmusaik hatékonyságát. Ezt a fejlesztési tevékenységet eddig is végezték, azonban a szabályozás keretét ad ahhoz, hogy az európai szabályozásra tekintettel vigyék végbe mindezt. Ennek égisze alatt, ha egy MI moderációs rendszer túl sok hibát vét (pl. jogszerűtlennek nem minősülő tartalmakat tilt le, vagy nem szűri ki a tiltott tartalmat), az a kockázatkezelési kötelezettségek megsértéséhez vezethet. Ilyen esetben a platform esetében nem az MI leállítása az elvárás, hanem a kockázat enyhítése – például az algoritmus továbbfejlesztése, szigorúbb emberi felülvizsgálat bevezetése vagy más korrekciós intézkedés. A Bizottság korai intézkedései is azt mutatják, hogy a hangsúly a megfelelő erőforrások és felügyelet biztosításán van. Erre mutat az az eljárás, amelyet a Bizottság 2023 végén indított az X platform ellen részben amiatt, hogy a cég drámaian csökkentette a moderátorok számát Európában (mintegy 20%-os létszámleépítés, 11 nyelvről 7-re szűkítve a lefedettséget). A Bizottság emiatt információkat kért az X-től arról, hogyan kívánja kezelni a tartalommoderációt és milyen kockázatelemzést végzett a generatív MI eszközök alkalmazásáról. Ez azt jelzi, hogy a Bizottságot aggodalommal tölti

⁵⁹ European Commission: Impact of the Digital Services Act on platforms. <https://tinyurl.com/4bb64t79>

⁶⁰ European Commission: Digital Services Act: Commission designates first set of very large online platforms and search engines. <https://tinyurl.com/y9ftz4ar>

⁶¹ Logan Partners: Digital Services Act (DSA) Compliance in Action: Insights from Leading Online Platforms. <https://tinyurl.com/ywww7b7fs>

⁶² European Commission: Very large online platforms and search engines publish first transparency reports under DSA. <https://tinyurl.com/4ftb5uhp>; Alessia Zornetta: Is the Digital Services Act Truly a Transparency Machine? The Tech Policy Press, Jul 11, 2024. <https://tinyurl.com/sbep8nwc>

el az elégtelen humán erőforrás és a felügyelet nélküli MI kombinációja. Ugyanakkor a TikTok esetében – amely növelte automatizált szűrésének arányát, de megtartotta a moderatori kapacitását – nem merült fel, hogy le kellene állítaniuk az MI-rendszereiket; inkább azt emelik ki, hogy a kettőt érdemes együtt alkalmazni a hatékony eredmény érdekében.⁶³ A gyakorlatban ez azt jelenti, hogy a platformok továbbra is támaszkodhatnak MI-megoldásokra a mérhető hatékonyság miatt, de kötelesek gondoskodni arról, hogy ezek az eszközök megbízhatóan működjenek, és hogy az automatizált döntéseket emberi felülvizsgálat és felhasználói jogorvoslat egészítse ki a DSA szabályai alapján.

A DSA ezek alapján jelentős előrelépést hozott a tartalommoderáció nyilvánossága és átláthatósága terén, és alkalmazkodásra készítette a nagy platformokat. Erre tekintettel kifejezetten pozitív szabályozási fejlemény, hogy az elaprózódó tagállami jogalkalmazás helyett a Bizottság széles körű hatáskörrel rendelkezik az *ex ante* jellegű, így az átláthatósági, a kockázatértékelési és kockázatcsökkentési kötelezettségek, valamint a válságkezelési mechanizmusok felügyeletére. Ezek a hatáskörök biztosítják, hogy az online óriásplatformok és keresőszolgáltatók megfelelően végrehajtsák a DSA előírásait, és proaktívan kezeljék a rendszerszintű kockázatokat. A Bizottság felügyeleti tevékenysége, amely magában foglalja az információgyűjtést, vizsgálatokat, szankciókat és a folyamatos nyomon követést, kulcsszerepet játszik a digitális szolgáltatások szabályozásának hatékonyságában és az alapvető jogok védelmében. E tekintetben célszerű lehet arra törekedni, hogy a Bizottság a DSA 66. cikke alapján megindult eljárásokban lehetőség szerint a nem megfelelést megállapító határozat létrehozatala helyett arra törekedjen, hogy – amennyiben az eset összes körülménye engedi – a szolgáltató kötelezettségvállalásokat tegyen, mert ez a megoldás egyszerre biztosíthatja az átláthatóság fejlesztését, a DSA szabályainak való megfelelést, s egyúttal a szolgáltatók bevonását a döntésbe és a számukra legkevésbé hátrányos technológiai megvalósítás kiválasztását is. Az így elért önkéntes jogkövetés fontos lehet, mert bár komoly represszív erővel rendelkező bírságtételekkel dolgozhat a Bizottság, végső soron elérheti, hogy a tagállami hatóságok – így a digitális szolgáltatási koordinátor jelzése alapján a tagállami bíróság – korlátozzák, elérhetetlenné tegyék ideiglenesen a szolgáltatást egy adott tagállamban a DSA 82. cikk (1) bekezdésében és a DSA 51. cikk (3) bek. b) pontjában foglaltaknak megfelelően, mégis ezek az intézkedések olyan költség- és időigényesek, hogy célszerűbb lehet a megfelelés gyorsabb módját elérni, ha arra lehetőség van. Ez akár a minősítési határokat elérő, de kisebb szereplők megfelelését is elősegíthetné, mert lehetőséget adna egyedi feltételek megállapítására.

A Bizottság végrehajtó munkájában azt is fontos lehet fókuszba állítani, hogy a rendszeres jelentések mellett szabályozással, egységesített módszertani megoldásokkal ösztönözni kell a szolgáltatókat valós idejű vagy ahhoz közelítő tájékoztatási, átláthatósági rendszerek bevezetésére, amelyek képesek lennének az egyedi szolgáltatási folyamatban is átfogó információkat nyújtani a felhasználóknak annak érdekében, hogy valóban érdemi átláthatóságérzetük alakulhasson ki, s a szabályozás alapján is megvalósulhasson az átláthatóság materiális (azaz nem formális, pusztán dokumentumok közzétételében mért) megvalósítása.

⁶³ TikTok Newsroom: DSA Third Transparency Report. <https://tinyurl.com/3hpum6zs>

ALGORITMIKUS ELFOGULTSÁG MINT JOGI DILEMMA

Lendvai Gergely Ferenc*

1. Bevezetés

A tanulmány bevezetéseként három példán keresztül szemléltetjük az algoritmikus elfogultság¹ megjelenését. A jelenség egyik legfontosabb pillanata közel tíz évvel ez előttre datálható. A COMPAS (*Correctional Offender Management Profiling for Alternative Sanctions*²) algoritmus – amelyet az amerikai büntető igazságszolgáltatási rendszerben a mai napig használnak – egyik korai fejlesztése egy olyan mesterséges intelligencia (a továbbiakban: „MI”) alapú eszköz volt, amelyet arra terveztek, hogy „megjósolja” a vádlottak visszaesésének valószínűségét, és ezzel támpontot adjon bírói döntésekhez és adatalapú szempontokkal szolgáljon a mérlegelés során.³ A rendszer a maga nemében forradalmi volt; a szoftver fejlesztői ugyanis azt állították, hogy a COMPAS befolyásmentesen és objektíven, kizárólag statisztikai alapon, adatvezérelt módon mutat eredményeket. A valóság ezzel szemben kevésbé volt idealisztikus. A rendszert már egészen korán kritizálták szakértők és kutatók annak tévedései és olykor kiszámíthatatlan döntési mintázatai,⁴ a szoftver módszertani hiányosságai⁵ és

* PhD-hallgató, Pázmány Péter Katolikus Egyetem, Jog- és Államtudományi Kar.
ORCID: <https://orcid.org/0000-0003-3298-8087>

¹ A szövegben következetesen az algoritmikus elfogultság kifejezést használjuk, mivel tartalmilag ez fedi a legjobban az angol „*algorithmic bias*” jelentését.

² Magyarul talán „Büntetés-végrehajtási elkövetői menedzsment profilalkotás az alternatív szankciókhoz” lenne a megfelelő fordítás.

³ Christoph Engel – Lorenz Linhardt – Marcel Schubert: Code is law: how COMPAS affects the way the judiciary handles the risk of recidivism. *Artificial Intelligence and Law*, Vol. 33. (2024) 3–4. <https://doi.org/10.1007/s10506-024-09389-8>

⁴ Tim Brennan – William Dieterich – Beate Ehret: Evaluating the Predictive Validity of the Compas Risk and Needs Assessment System. *Criminal Justice and Behavior*, Vol. 36., No. 1. (2008) <https://doi.org/10.1177/0093854808326545>

⁵ Jay P. Singh: Predictive Validity Performance Indicators in Violence Risk Assessment: A Methodological Primer. *Behavioral Sciences & the Law*, Vol. 31., No. 1. (2013) 8–12. <https://doi.org/10.1002/bsl.2052>



etikai dilemmái miatt.⁶ 2016-ban viszont az algoritmikus elfogultság rövid, de igen gazdag történetének egyik legjelentősebb ügyével került a köztudatba a COMPAS. A ProPublica független, non-profit, főleg oknyomozó újságírára szakosodott szervezet széleskörű vizsgálatot folytatott a szoftver által javasolt döntések kapcsán és azt állította, hogy az algoritmus jelentős faji elfogultságot mutat az afro-amerikai emberek hátrányára.⁷ A szervezet kutatói és újságírói a floridai Broward megyében több mint 10000 vádlott adatait vizsgálta meg, kifejezetten a lehetséges faji elfogultságra összpontosítva. A Broward megyei seriff hivatalától két év COMPAS-pontszámát szerezték meg, amely a 2013-ban és 2014-ben „pontozott” összes személyt felöleli. A hatalmas adatmennyiséget magában foglaló vizsgálat végén a ProPublica kiszámította az algoritmus általános pontosságát, és feltárta, hogy a nem visszaeső afro-amerikai vádlottakat majdnem kétszer nagyobb valószínűséggel minősítették tévesen magas kockázatúnak, mint fehér társaikat (45% vs. 23%), és a visszaeső fehér vádlottakat gyakrabban minősítették tévesen alacsony kockázatúnak, mint az afro-amerikai visszaesőket (48% vs. 28%). Az erőszakos visszaesés-elemzés még drasztikusabb képet vázolt fel: a korábbi bűncselekmények, a jövőbeni visszaesés, az életkor és a nem figyelembevételével a COMPAS 77%-kal nagyobb valószínűséggel adott magasabb kockázati pontszámot az afro-amerikai vádlottaknak, mint a fehér vádlottaknak. Bár a ProPublica elemzését Flores és társai módszertani hiányosságokra és hibás adatokra tekintettel sok helyütt megkérdőjelezték,⁸ a COMPAS és az algoritmus faji alapú megkülönböztetésre okot adó gyakorlatáról a mai napig születnek empirikus kutatások, javarészt megerősítve a ProPublica állításait.⁹

A második, talán a mai napig az egyik leghíresebb esete az algoritmikus elfogultságnak az Amazon MI-alapú toborzó alkalmazása és rövid „élete”. A BBC által „szexista MI”-nek¹⁰ titulált szoftvert 2014-ben kezdte el fejleszteni az Amazon azzal

⁶ Georgios Bouchagiar: Is Europe prepared for Risk Assessment Technologies in criminal justice? Lessons from the US experience. *New Journal of European Criminal Law*, Vol. 15., No. 1. (2024) <https://doi.org/10.1177/20322844241228676>

⁷ Julian Angwin – Jeff Larson – Surya Mattu – Lauren Kirchner: Machine Bias. *ProPublica*, 2016. május 23. <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing> A tanulmányhoz használt módszertanhoz ld. Jeff Larson – Surya Mattu – Lauren Kirchner – Julia Angwin: How We Analyzed the COMPAS Recidivism Algorithm. *ProPublica*, 2016. május 23. <https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm>

⁸ Anthony W. Flores – Kristin Bechtel – Christopher T. Lowenkamp: False Positives, False Negatives, and False Analyses: A Rejoinder to ‘Machine Bias: There’s Software Used across the Country to Predict Future Criminals and It’s Biased against Blacks.’ *Federal Probation*, Vol. 80., No. 2. (2016) 44–45. Ugyancsak ld. Cynthia Rudin – Caroline Wang – Beau Coker: The Age of Secrecy and Unfairness in Recidivism Prediction. *Harvard Data Science Review*, Vol. 2., No. 1. (2020) <https://doi.org/10.1162/99608f92.6ed64b30>

⁹ Riccardo Fogliato – Alexandra Chouldechova, – Max G’Sell: Fairness Evaluation in Presence of Biased Noisy Labels. *International Conference on Artificial Intelligence and Statistics*, 2020. június 3. (2020) 2325–2336.; Francesca Lagioia – Riccardo Rovatti – Giovanni Sartor: Algorithmic fairness through group parities? The case of COMPAS-SAPMOC. *AI & Society*, Vol. 38., No. 2. (2022) 463–464. <https://doi.org/10.1007/s00146-022-01441-y>

¹⁰ Amazon scrapped ‘sexist AI’ tool. *BBC*, 2018. október 10. <https://www.bbc.com/news/technology-45809919>

a céllal, hogy a leendő alkalmazottak felvételi folyamatát egyszerűsítsék. Maga a mechanizmus tipikusan olyan feladatokat érintett, amelyek a humánerőforrás területén dolgozók munkakörét fedik le: önéletrajzokat elemzett és szortírozott, illetve a különböző pozíciókhoz szükséges legjobb jelentkezőket igyekezett azonosítani az önéletrajzban feltüntetett információk alapján. Az eszközt sajátos módon az Amazon a saját, egy évtized alatt a vállalathoz benyújtott önéletrajzokon képezte ki, és gépi tanulási algoritmusokat használtak a minták azonosítására és a jelöltek értékelésére. A program egy specifikus, egytől öt csillagig terjedő pontozást vezetett be.¹¹ 2018-ra nyilvánvalóvá vált, hogy az eszköz szisztematikusan diszkriminálja a nőket, különösen a műszaki szerepkörökben, amely jelentős hibát fedett fel a mögöttes adatokban és algoritmusokban.¹² Az előítéletesség gyökere a képzési adatokban rejlett: a korábbi önéletrajzok ugyanis túlnyomórészt férfi jelentkezőktől származtak és az algoritmus megtanulta, hogy a sikeres, felvett és alkalmazandó jelöltekhez a férfiakat társítsa.¹³

Végül egy egészen friss, nagy médiavisszhangot generáló eseménnyel érdemes zárni a példák sorát. Az MI eddig soha nem látott demokratizálódása során, amelyben jelentős szerepet játszottak a különböző nagy nyelvi modellek elterjedése (például ChatGPT), heves verseny alakult ki a képgeneráló MI-rendszerek fejlesztése vonatkozásában, amelynek 2024-ben a legprominensebb résztvevői a DALL-E (ma már a ChatGPT-be integrált) és a MidJourney nevű modellek voltak. Új versenyzőként a Google kifejlesztette a Gemini (Bard) névre hallgató rendszerét, amelynek egyik funkciója a képgenerálás volt.¹⁴ Ez utóbbi funkció bemutatkozása ugyanakkor igen ellentmondásosnak bizonyult. Ellentétben ugyanis a fenti két példával, amelyek esetében az az adatkészletekben rejlő „inherens” társadalmi alapú megkülönböztetés volt a polémiák alapja, a Gemini esetében pont az ilyen jellegű problémák kiszűrésére irányuló „túlkompenzálás” volt az elfogultság jelenségének eredője. A Gemini ugyanis többször is történelmileg pontatlan és etnikailag sokszínű ábrázolásokat készített fehér vagy hagyományosan fehérnek ábrázolt emberekről és történelmi szereplőkről. Ennek két eklatáns példája az egészen bizarr hatást keltő afro-amerikai nemszocialista katona (bal) és az amerikai indián (*amerind*) viking harcos (jobb). (1. ábra)

¹¹ Az Amazonon árult termékek értékeléséhez volt hasonló a pontozási rendszer.

¹² Lennart Hofeditz – Milad Mirbabaie – Audrey Luther – Riccarda Mauth – Ina Rentemeister: Ethics Guidelines for Using AI-based Algorithms in Recruiting: Learnings from a Systematic Literature Review. *Proceedings of the ... Annual Hawaii International Conference on System Sciences/Proceedings of the Annual Hawaii International Conference on System Sciences*, 2022. január 1. (2022) 145. <https://doi.org/10.24251/hicss.2022.018>

¹³ Jeremy Hsu: Can AI hiring systems be made antiracist? Makers and users of AI-assisted recruiting software reexamine the tools' development and how they're used. *IEEE Spectrum*, Vol. 57., No. 9. (2020) 9. <https://doi.org/10.1109/mspec.2020.9173891>

¹⁴ Vö. Muhammad Imran – Norah Almusharraf: Google Gemini as a next generation AI educational tool: a review of emerging educational technology. *Smart Learning Environments*, Vol. 11., No. 1. (2024) <https://doi.org/10.1186/s40561-024-00310-z> ; Hamid Reza Saeidnia: Welcome to the Gemini era: Google DeepMind and the information industry. *Library Hi Tech News*, December 26, 2023. <https://doi.org/10.1108/lhtn-12-2023-0214>



1. ábra. A Gemini által generált képek (Forrás: Matthew Field, *Telegraph*, 2024¹⁵)

A Gemini kimeneteinek elfogultsága a korábbi két esethez hasonlóan az MI képzéséhez használt adatokban gyökerezik, amelyek bizonyos nézőpontokat és kérdéseket, mint például az inkluzivitást felül-, míg másokat, így a történelmileg hű ábrázolást alulreprezentáltak, ami bár nem szándékolt, de kétséget kizárólag egyaránt szürreális és súlyosan sértő eredményekhez vezettek. Bár a Google azonnal reagált és elnézést kért, illetve rendkívüli gyorsasággal kigyomlálta a problémára okot adó elemeket,¹⁶ a Gemini-t a mai napig kritizálják etnikai elfogultsága miatt.¹⁷

E három példa hűen tükrözi Cathy O’Neil, az algoritmikus elfogultság egyik legelismertebb kutatójának gondolatát, aki az elsők között fejtette ki, hogy az új technológiákkal nem megszűntetjük, hanem csupán álcázzuk az emberi elfogultságot.¹⁸ Bár az adatok kétségtelenül, természetükből adódóan függetlennek és objektíveknek tűnhetnek, az adatokkal dolgozó egyén viszont nem az – éppen ebből ered az elfogultság. Jelen tanulmány célja, hogy az algoritmikus elfogultságot konceptualizálja és a jelenlegi jogi válaszokat bemutassa. E körben elsősorban az európai jogalkotásra fókuszálunk

¹⁵ Matthew Field: From Black Nazis to female Popes and American Indian Vikings: How AI went ‘woke’. *The Telegraph*, 2024. február 23. <https://tinyurl.com/2uke3bt9>

¹⁶ Ali Robertson: Google apologizes for ‘missing the mark’ after Gemini generated racially diverse Nazis. *The Verge*, 2024. február 21. <https://tinyurl.com/mfsvs2xc>

¹⁷ Julia Barroso Da Silveira – Ellen Alves Lima: Racial Biases in AIs and Gemini’s Inability to Write Narratives About Black People. *Emerging Media*, Vol. 2., No. 2. (2024) <https://doi.org/10.1177/27523543241277564>

¹⁸ Cathy O’Neil: *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. London, Penguin, 2016. 19.

és igyekszünk feltárni olyan hiányosságokat, amelyeket érdemes a jövő jogalkotóinak orvosolni. A tanulmány jogi elemzésre és szakirodalmi áttekintésre épül. Bár az algoritmikus elfogultság hazai, magyar nyelvű irodalma igen kezdetleges, reméljük, hogy jelen dolgozat vitaindító és inspiráló lehet a digitális technológiák és a mesterséges intelligencia szabályozása iránt érdeklődő kutatók számára.

2. Az algoritmikus elfogultság fogalma

Akárcsak a digitális jelenségek javarészában, az algoritmikus elfogultság esetében is problémás az egységes fogalom megalkotása.¹⁹ Általánosságban véve az algoritmikus elfogultság alatt olyan MI-rendszerekben vagy MI-alapú rendszerekben előforduló szisztematikus és strukturált hibákat és elfogultsági pontokat értünk, amelyek részrehabjló²⁰ eredményeket és egyenlőtlenségeket eredményeznek anélkül, hogy erre valamilyen módon igazolható indok lenne.²¹ Javasolt a fogalmat elemeire bontani. Az algoritmikus elfogultság és az MI kapcsolata talán magától értetődőnek hat, ugyanakkor érdemes röviden kitérni ennek a hátterére. Ahogy Shin és Shin is aláhúzta, az emberi kognitív előítéletek sokszor „bekerülnek” az algoritmusokba, az MI pedig képes felerősíteni ezeket.²² Ahogy a szerzők fogalmaznak: „ahhoz, hogy az MI igazodjon az emberek által kívánatosnak tartott funkciókhoz és feladatokhoz, meg kell tanulnia ezeket a preferenciákat, viszont az emberi értékek megtanulása kockázatokat hordoz magában.”²³ Az elfogultság és az MI kapcsolata tehát technológiai szempontból inherens.²⁴ Az algoritmikus elfogultság fogalmának központi eleme a szisztematikus, rendszerszintű elfogultság. Bár a rendszerszintűség és az elfogultság irodalma igen gazdag,²⁵ jelen kontextusban két definíciós elemre érdemes kitérni. Egyfelől, a szisztematikus jelleg alatt nem egyedi, egyszeri hibából adódó esetekre gondolunk.

¹⁹ Jin-Young Kim – Sung-Bae Cho: An information theoretic approach to reducing algorithmic bias for machine learning. *Neurocomputing*, Vol. 500. (2022) 26–27. <https://doi.org/10.1016/j.neucom.2021.09.081>

²⁰ A szakirodalom az angol „unfair” szóval él, ugyanakkor az „igazságtalan” fordítás fogalmilag nem érintené az algoritmikus elfogultság elfogultságra, vagy akár bizonyos értelemben méltánytalanságra utaló jellegét. A fogalmi kérdésekhez lásd a „fairness” és „justice” distinkciót itt: Lionel P. Robert – Casey Pierce – Liz Marquis – Sangmi Kim – Rasha Alahmad: Designing fair AI for managing employees in organizations: a review, critique, and design agenda. *Human-Computer Interaction*, Vol. 35., No. 5–6. (2020) <https://doi.org/10.1080/07370024.2020.1735391>

²¹ Nima Kordzadeh – Maryam Ghasemaghahi: Algorithmic bias: review, synthesis, and future research directions. *European Journal of Information Systems*, Vol. 31., No. 3. (2021) 388. <https://doi.org/10.1080/0960085x.2021.1927212>

²² Donghee Shin – Emily Y. Shin: Data’s Impact on Algorithmic Bias. *Computer*, Vol. 56., No. 6. (2023) 90. <https://doi.org/10.1109/mc.2023.3262909>

²³ Uo.

²⁴ Michael Veale – Max Van Kleek – Reuben Binns: Fairness and Accountability Design Needs for Algorithmic Support in High-Stakes Public Sector Decision-Making. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, Montreal, QC, Canada*, February 4, 2018. (2018) <https://doi.org/10.31235/osf.io/8kvf4>

²⁵ Ehhez ld. Gabbrielle M. Johnson: Algorithmic bias: on the implicit biases of social technology. *Synthese*, Vol. 198., No. 10. (2020) <https://doi.org/10.1007/s11229-020-02696-y>

Visszatérve a Gemini példájára, a képgeneráló rendszerek gyakran tévesztenek, vagy állítanak elő nem megfelelő kimenetet, a Google fejlesztése esetében viszont egy folyamatosan, visszatérő módon előforduló problémáról volt szó, amely nem egy egyedi „glitch”, hanem – *ab ovo* – a modell kialakítottságának sajátja volt. Másfelől, a rendszerszintűség egyet jelent a reprodukálhatósággal is,²⁶ amely leginkább az adatgyűjtés és -feldolgozás során ejtett hibákból és a megfelelő transzparencia-mechanizmusok hiányából,²⁷ illetve – éppen a Gemini esetében – az adatkészletben rejlő „adott” hibák „túlbuzgó” emberi tevékenységből adódik. Talán a legnehezebben megragadható fogalmi elempár a részrehajló eredmények és egyenlőtlenségek, mint kimenetek. Ennek oka sokrétű. Egyrésről a részrehajló „jelleg” és az egyenlőtlenség korántsem technológiai fogalmak. Ahogy Kordzadeh és Ghasemaghahi összefoglaló tanulmányából is kiderül, a részrehajlásnak és „elfogultságnak” igen gazdag „techno-filozófiai” szakirodalma van.²⁸ Ha bölcséleti oldalról vizsgáljuk a részrehajlás kérdését, akkor a szerzők szintézisére hivatkozva az elfogultság a következő jelenségeket takarja:

1. egy algoritmus egyenlőtlenül osztja el az előnyöket és hátrányokat a különböző egyének vagy csoportok között;
2. ez az egyenlőtlen elosztás az egyének eredendő tulajdonságaiban, tehetségében vagy szerencséjében rejlő különbségekre vezethető vissza, így
3. az algoritmikus elfogultság egy algoritmus kimeneteiben megjelenő, az egyenlőségtől való szisztematikus eltérés.²⁹

Másrésről az egyenlőtlenség mint probléma lényegében implikálja, hogy az algoritmusok kialakításakor létezik egy etikai standard, amelynek alapja, hogy az algoritmus alapú „egyenlőség” elérhető és megvalósítható, ha a nem szándékosan beépített torzításokat és részrehajlásokat kiszűrjük. Ezt egyes kutatók „észlelt pártatlanságnak” (*perceived fairness*) nevezik, amely utal az algoritmusok által generált kimenetek, a döntéshozatal és az algoritmusok kialakításának pártatlan, azaz nem megkülönböztető jellegére.³⁰

²⁶ Zhisheng Chen: Ethics and discrimination in artificial intelligence-enabled recruitment practices. *Humanities and Social Sciences Communications*, Vol. 10., No. 1. (2023) 3. <https://doi.org/10.1057/s41599-023-02079-x>

²⁷ Jonathan Dodge – Q. Vera Liao – Yunfeng Zhang – Rachel K. E. Bellamy – Casey Dugan: Explaining models: an empirical study of how explanations impact fairness judgment. *Proceedings of the 24th International Conference on Intelligent User Interfaces. Marina Del Ray, CA, USA.*, March 17, 2019. (2019) 275–285. <https://doi.org/10.1145/3301275.3302310>

²⁸ Kordzadeh–Ghasemaghahi i. m. 7–8.

²⁹ Uo.; Sahil Verma – Julia Rubin: Fairness definitions explained. *2018 IEEE/ACM International Workshop on Software Fairness (FairWare), Gothenburg, Sweden.*, May 29, 2018. (2018) <https://doi.org/10.1145/3194770.3194776>; Reuben Binns Max Van Kleek – Michael Veale – Ulrik Lyngs – Jun Zhao – Nigel Shadbolt: ‘It’s Reducing a Human Being to a Percentage’: Perceptions of Justice in Algorithmic Decisions. *arXiv (Cornell University)*, April 21, 2018. (2018) 377. <http://arxiv.org/abs/1801.10408>

³⁰ Binns et al. i. m.; Kordzadeh–Ghasemaghahi i. m. 9–10.

3. Az algoritmikus elfogultság és a szabályozás kapcsolata, illetve az amerikai perspektívák

A jogi szabályozás az algoritmikus elfogultság esetében polemikus: kérdés tucatjával érkezik a szabályozó felé, a válaszokra viszont sokszor éveket kell várni, sőt, ha mégis sikerül megoldást találni, talán már okafogyottá is válik az eredeti kérdés.³¹ Jelen szegmensben az algoritmikus elfogultság szabályozási dilemmáit mutatjuk be, azaz, hogy egyáltalán *miért* jogi dilemma az algoritmikus elfogultság kezelése.

Az algoritmikus elfogultság elsősorban adatvédelmi és a mesterséges intelligencia általános szabályozását érinti. A „szabályozás” e körben nem specifikus, kifejezetten az elfogultságra vonatkozó jogszabályokat, hanem bizonyos nagyobb tárgykörök vonatkozó szabályait érinti, amelyek alkalmazhatók az algoritmikus elfogultságra. Nem túlzás azt állítani, hogy a jelenséget érintő szabályozási keretet a nemzeti, regionális és globális szintű kezdeményezések sokasága jellemzi, nem is beszélve a platformok és a mesterséges intelligenciával foglalkozó cégek saját szabályairól. Fejlődő, az MI-szabályozás alszegmensének tekinthető szabályozási környezetről beszélhetünk tehát, amelyet leginkább a harmonizáció alacsony szintje és a töredezettség jellemez.³² Ahogy más mesterséges intelligenciát érintő kérdésekben,³³ az algoritmikus elfogultság esetén is alapvetően két fő szabályozási irány azonosítható: az európai, erősen korlátozó, ugyanakkor felhasználó-központú és az amerikai, fragmentáltabb, „kevert” jellegű, a technológiai és gazdasági fejlődést is támogató attitűd.³⁴ Tekintve, hogy az európai szabályozást a következő fejezetben részletezem, a következőkben az amerikai megközelítést ismertetem röviden.

Wang és társai szerint az amerikai jogi keretrendszer az algoritmikus elfogultság tekintetében az alapvető polgárjogi védelemben és a tizennegyedik módosításban gyökerezik, kiemelt hangsúllyal három alapvető elvre: az egyenlőségre, a megkülönböztetésmentességre és az átláthatóságra.³⁵ A fentebb is említett „kevert” szabályozási rendszer abból fakad, hogy az algoritmikus elfogultságra egyaránt irányadók a vonatkozó, elsősorban munkakereséssel, hitelfelvétellel és egy állásbetöltéssel kapcsolatos törvények, mint a *Fair Credit Reporting Act* és az *Equal Employment Opportunity Commission*, amelyek többek között kitérnek az algoritmusok pártatlanságának fontosságára is,³⁶ ezen túlmenően pedig a bírósági ítéletek is jelentős szere-

³¹ Gergely Ferenc Lendvai: Sharenting as a regulatory paradox – a comprehensive overview of the conceptualization and regulation of sharenting. *International Journal of Law Policy and the Family*, Vol. 38., No. 1. (2024) 17. <https://doi.org/10.1093/lawfam/ebae013>

³² Xukang Wang – Ying Cheg Wu – Xueliang Ji – Hongpeng Fu: Algorithmic discrimination: examining its types and regulatory measures with emphasis on US legal practices. *Frontiers in Artificial Intelligence*, Vol. 7. (2024) 4–6. <https://doi.org/10.3389/frai.2024.1320277>

³³ Vö. Gosztonyi Gergely – Lendvai Gergely: Deepfake és dezinformáció. Mit tehet a jog a mélyhamisítással készített álhírek ellen? *Médiakutató*, Vol. 25., No. 1. (2024) <https://doi.org/10.55395/mk.2024.1.3>

³⁴ Wang et al. i. m. 4–6

³⁵ Uo.

³⁶ Mark MacCarthy: Standards of Fairness for Disparate Impact Assessment of Big Data Algorithms. *Cumberland Law Review*, Vol. 48., No. 1. (2017) 74–89.

pet játszanak mind az e törvények értelmezésében az algoritmikus diszkriminációval kapcsolatos esetekben, mind az egyre gyakrabban előforduló, automatizált rendszerek által okozott foglalkoztatási és lakhatási előítéltelességgel foglalkozó ügyekben.³⁷

Átfogó szövetségi törvény, amely kifejezetten az algoritmikus elfogultság kezelésére szolgálna viszont nincs, még annak ellenére sem, hogy az e kézirat írásakor leköszönő Joe Biden elnök 2023 októberében kiadott végrehajtási rendeletében (*executive order*) a biztonságos, megbízható mesterséges intelligenciára vonatkozóan külön kiemelésre került, hogy az amerikai szabályozás kifejezetten ügyelni fog az algoritmikus elfogultság megfelelő kezelésére.³⁸ A dokumentum erre több javaslatot is tesz:

1. Egyértelmű iránymutatást kell adni a bérbeadók, a szövetségi juttatási programok és a szövetségi vállalkozók számára annak megakadályozása érdekében, hogy a mesterséges intelligencia algoritmusokat a megkülönböztetés súlyosbítására használják.
2. Az Igazságügyi Minisztérium és a szövetségi polgárjogi hivatalok közötti együttműködést biztosítani és támogatni kell, hogy az MI-vel és kifejezetten az algoritmikus elfogultsággal érintett polgári jogi jogsértéseket megfelelően ki lehessen vizsgálni.
3. A büntető igazságszolgáltatási rendszerben biztosítani kell az algoritmikus pártatlanságot és méltányosságot.
4. A 2) és 3) pontok esetében ki kell alakítani „legjobb gyakorlatokat”.³⁹

Bár a Biden-adminisztráció ambiciózusnak tűnt 2023 őszén, amelyet megelőzőtt egy évvel korábban a „*Blueprint for an AI Bill of Rights*” javaslat,⁴⁰ az amerikai MI-szabályozás jövője a Trump-adminisztráció alatt teljes fordulatot vett. Lényegében pár órával beiktatása után ugyanis a republikánus elnök visszavonta a Biden-féle végrehajtási rendeletet, sőt, a Fehér Ház kommunikációjában „ártalmasnak” nyilvánította azt,⁴¹ egy későbbi kiadványban pedig kifejezetten aláhúzta, hogy semmiféle „fejlesztés-korlátozó” szabályozást nem kíván bevezetni a mesterséges intelligencia tekintetében.⁴² Ahogy azt Saed is kifejti, a Trump-adminisztráció radikálisan liberális MI-(nem) szabályozásának következményei hosszú-, de még rövidtávon is beláthatatlanok, többek között amiatt is, mert az MI-fejlesztéssel járó kockázatok felmérésére nem került

³⁷ Wang et al. i. m. 4.

³⁸ Fact Sheet: President Biden Issues Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence. *WH.Gov*, 2023. október 30. (A kézirat leadásakor elektronikusan elérhető volt, de a Trump-adminisztráció azóta törölte.)

³⁹ Uo.

⁴⁰ Blueprint for an AI Bill of Rights. *WH.Gov*, 2022–2023. (A kézirat leadásakor elektronikusan elérhető volt, de a Trump-adminisztráció azóta törölte.)

⁴¹ Initial Rescissions of Harmful Executive Orders and Actions. White House, 2025. január 20. <https://tinyurl.com/4wuhbe8h>

⁴² Removing Barriers to American Leadership in Artificial Intelligence. White House, 2025. január 23. <https://tinyurl.com/2u6zf73x>

különösebb hangsúly eddig.⁴³ Azt viszont egészen biztosan állíthatjuk, hogy a Biden-adminisztráció által felvázolt progresszív, az emberek biztonságát előtérbe helyező MI-szabályozás visszatérte belátható időn belül nem valószínű. Bár a kutatások e téren egyelőre spekulatívak, Novelli és társainak tanulmánya fontos megállapításokat tesz. A szerzők két lehetséges irányt vázolnak fel az AI-szabályozás terén a Trump-adminisztráció alatt: az egyik a decentralizált, állami szintű szabályozás, ahol az egyes államok saját törvényeket alkotnak, a másik pedig a „szövetségi szabályozás térnyerése”, amelynek célja az egységes, dereguláción alapuló, innovációbarát szabályozási környezet megteremtése.⁴⁴ Mindkét modell viszont komoly jogi és politikai kihívásokat vet fel, legyen szó a fragmentált szabályozás kockázatáról, vagy éppen a potenciális alkotmányjogi konfliktusokról.

4. Az uniós jog és az algoritmikus elfogultság ezer arca

Az algoritmikus elfogultság kapcsán – jelen dolgozat írásakor – egyértelműen az európai jogi keretrendszer nyújtja a legelőremutatóbb szabályozási megközelítést, még akkor is, ha egyesek hevesen kritizálják az uniós jogot amiatt, mert nem tér ki minden potenciális kockázatra.⁴⁵ Több oldalról érdemes megközelíteni az uniós szabályozást, a következőkben specifikus esetkörökre bontva ismertetjük a jogszabályok adta lehetőségeket.

Először érdemes az általános, ugyanakkor kifejezetten algoritmus-specifikációtól mentes keretrendszert ismertetni. Ehhez elsőként az adatvédelmi kérdéseket és az Általános adatvédelmi rendelet (GDPR)⁴⁶ megoldásait mutatjuk be. A GDPR 22. cikke e körben kulcsfontosságú. E cikk (1) bekezdése szerint ugyanis az „érintett jogosult arra, hogy ne terjedjen ki rá az olyan, kizárólag automatizált adatkezelésen – ideértve a profilalkotást is – alapuló döntés hatálya, amely rá nézve joghatással járna vagy őt hasonlóképpen jelentős mértékben érintené.”⁴⁷ E rendelkezés alapja az a preambulumban is megfogalmazott cél, hogy a személyes adatok kezelése az embereket kell, hogy szolgálja,⁴⁸ illetve hogy biztosított legyen a tisztességes, és legfőképp átlátható adatkezelés.⁴⁹ A rendelkezés fő kérdése, különösen az algoritmikus elfogultság kapcsán

⁴³ Ferial Saed: The Uncertain Future of AI Regulation in a Second Trump Term. *Stimson*, 2025. március 13. <https://tinyurl.com/bnuzdttt>

⁴⁴ Claudio Novelli – Akriti Gaur – Luciano Floridi: Two Futures of AI Regulation under the Trump Administration. (Preprint.) *SSRN*, 2025. március 30. <http://dx.doi.org/10.2139/ssrn.5198926>

⁴⁵ Ehhez ld. Philipp Hacker: Teaching fairness to artificial intelligence: Existing and novel strategies against algorithmic discrimination under EU law. *Common Market Law Review*, Vol. 55., No. 4. (2018) 1143–1146. <https://doi.org/10.54648/cola2018095>

⁴⁶ Az Európai Parlament és a Tanács (EU) 2016/679 rendelete (2016. április 27.) a természetes személyeknek a személyes adatok kezelése tekintetében történő védelméről és az ilyen adatok szabad áramlásáról, valamint a 95/46/EK rendelet hatályon kívül helyezéséről (általános adatvédelmi rendelet) (EGT-vonatkozású szöveg) HL L 119., 04/05/2016,1–88. o.

⁴⁷ GDPR 22. cikk (1) bekezdés.

⁴⁸ GDPR 4. preambulumbekzdés.

⁴⁹ GDPR 60. preambulumbekzdés.

a joghatás, illetve a „hasonlóképpen jelentős mértékű” más fajta döntés definiálása, amelyre a GDPR csak absztraháltan vállalkozik a preambulumban.⁵⁰ A joghatás alatt elsősorban olyan döntéseket vagy tevékenységeket értünk, amely érinti valakinek a jogait, jogi státuszát, vagy akár szerződéses jogait. Barbosa és Félix szerint ilyenre lehet példa a szavazati jogra, a fogyatékoság miatti havi nyugdíjra való jogosultságra, vagy egy országba való belépési képességre gyakorolt hatások.⁵¹ A hasonlóképpen jelentős mértékű hatással járó döntés interpretálása egy fokkal viszont komplexebb.⁵² Bár a 71. preambulumbekzdés ad némi támaszt,⁵³ példálózó eseteken túlmenően nem derül ki a GDPR szövegéből mi a pontos köre az ilyen típusú döntéseknek.⁵⁴ Barbosa és Félix e helyütt az javasolja, hogy a rendelkezést elsősorban annak súlya, mintsem automatikus jellege szerint kell értékelni; azaz, bár kétségtelenül automatizált módon ajánl egy zeneszámot a streaming-platform, feltételezhetően ez nem jár szignifikáns, vagy akár joghatással összehasonlítható hatással.⁵⁵ Fontos megjegyezni viszont, hogy vannak kivételek a 22. cikk (1) bekezdése alól, így ha *A*) szerződéskötés vagy szerződés teljesítés miatt szükséges az adatkezelés, *B*) az uniós jog lehetőséget ad rá, illetve ha *C*) az érintett kifejezetten hozzájárul, nem szükséges a fenti rendelkezést alkalmazni.⁵⁶ Az *A*) és *C*) esetkörnél viszont fontos kitétel említi meg a Rendelet, mivel „az adatkezelő köteles megfelelő intézkedéseket tenni az érintett jogainak, szabadságainak és jogos érdekeinek védelme érdekében, ideértve az érintettnek legalább azt a jogát, hogy az adatkezelő részéről emberi beavatkozást kérjen, álláspontját kifejezze, és a döntéssel szemben kifogást nyújtson be.”⁵⁷ Ahogy Bygrave is jelzi, ezek a lehetőségek ugyan alapvetően *ex post* jellegűek, mégis rendkívül előremutatóak a felhasználó- és az adatvédelem területén, és jelentős garanciákat hordoznak egy potenciális, sértő algoritmikus elfogultság eredményeként megszületett döntéssel szemben.⁵⁸

⁵⁰ GDPR 70. preambulumbekzdés.

⁵¹ Sandra Barbosa – Sara Félix: Algorithms and the GDPR: An analysis of article 22. *Anuário da Proteção de Dados*, 2021. 75.

⁵² Uo. 75–76.

⁵³ [...] természetes személyekre vonatkozó személyes jellemzők bármilyen automatizált személyes adatok kezelése keretében történő kiértékelése, különösen az érintett munkahelyi teljesítményére, gazdasági helyzetére, egészségi állapotára, személyes preferenciáira vagy érdeklődési körökre, megbízhatóságra vagy viselkedésre, tartózkodási helyére vagy mozgására vonatkozó jellemzők elemzésére és előrejelzésére.

⁵⁴ Lee A. Bygrave: Article 22. In: Christopher Kuner –Lee A Bygrave – Christopher Docksey – Laura Drechsler (szerk.): *The EU General Data Protection Regulation (GDPR): A Commentary*. Oxford, Oxford University Press, 536.

⁵⁵ Barbosa–Félix i. m. 76.

⁵⁶ GDPR 22. cikk (2) bekezdés.

⁵⁷ GDPR 22. cikk (3) bekezdés.

⁵⁸ Bygrave i. m. 538.

Bár a DSA, azaz a digitális szolgáltatásokról szóló rendelet⁵⁹ elsősorban a platform-szabályozás területén vezet be úttörő szabályokat,⁶⁰ részben lehet alkalmazni a rendelkezéseit az algoritmikus elfogultságra is. A DSA – Husovec szavaival élve – „megkülönböztető” szabályozást alkalmaz, amelynek értelmében a különböző platformok azok kategóriája és mérete alapján vezet be szabályokat.⁶¹ Az algoritmusok kérdése két helyütt kerül elő hangsúlyosan a DSA-ban. Elsőként a 14. cikkben kerül említésre, amely a szerződési feltételekről rendelkezik, és külön kitér arra, hogy az algoritmikus döntéshozatalról szóló információkat a közvetítő szolgáltatóknak „világosan, egyszerűen, érthetően, felhasználóbarát módon” kell megfogalmazniuk és könnyen hozzáférhetővé kell tenniük az adott szolgáltatást igénybe vevő számára.⁶² Jelentősebb fókusz kerül az algoritmusokra és azok átláthatóságára az online óriásplatformok és nagyon népszerű online keresőprogramok esetében.⁶³ Ezek olyan platformok, amelyek „havonta átlagosan legalább 45 millió, a szolgáltatást aktívan igénybe vevővel rendelkeznek az Unióban.”⁶⁴ E helyütt megjegyzendő viszont, hogy konjunktív feltétele az online óriásplatform státuszuk az is, hogy az Európai Unió Bizottsága annak is minősítse az adott platformot.⁶⁵ A minősítési eljárás nem egyszeri, és a Bizottság folyamatosan ellenőrzi és bővíti a VLOPSE-k listáját,⁶⁶ hogy a lehető legteljesebb körű képet lehessen kapni arról, kik is a platformok „óriásai”. A VLOPSE-kra számos új, igen progresszív szabály vonatkozik; az algoritmusok kérdéskörét tekintve viszont a kockázatértékelés kötelezettsége a legszignifikánsabb új szabály. A DSA 34. cikk értelmében ugyanis a VLOPSE-knak „gondosan azonosítaniuk, elemezniük és értékelniük kell a szolgáltatásaik és a kapcsolódó rendszerek – többek között az algoritmikus rendszerek – kialakításából vagy működéséből vagy a szolgáltatásaik használatából eredő rendszerszintű kockázatokat”.⁶⁷ Ilyen rendszerszintű kockázatok közé tartozik többek között a jogellenes tartalmak megosztása, az alapvető jogokat érő negatív hatások és gyakorlatok, a demokratikus diskurzust és a közbiztonságot érintő veszélyek, illetve nemi alapú erőszakkal, a közegészség és a kiskorúak védelmével összefüggő, továbbá a személyek testi és szellemi jóllétére gyakorolt negatív következmények.⁶⁸ Az algoritmikus elfogultság e rendszerszintű kockázatok mindegyik formájában manifesztálódhat – gon-

⁵⁹ Az Európai Parlament és a Tanács (EU) 2022/2065 rendelete (2022. október 19.) a digitális szolgáltatások egységes piacáról és a 2000/31/EK irányelv módosításáról (digitális szolgáltatásokról szóló rendelet) (EGT-vonatkozású szöveg). PE/30/2022/REV/1; HL L 277., 2022.10.27, 1–102. o.

⁶⁰ Gergely Ferenc Lendvai: Taming the Titans? – Digital Constitutionalism and the Digital Services Act. *ESSACHESS*, Vol. 17., No. 34. (2024) 169–172. <http://dx.doi.org/10.21409/0M3P-A614>

⁶¹ Martin Husovec: The DSA’s Scope Briefly Explained. *SSRN*, July 4, 2023. 1–2.

⁶² Ehhez még ld. DSA 45. preambulumbekzdés.

⁶³ E két platformfajtát legtöbbször „VLOPSE” néven rövidítjük, utalva az angol elnevezésre (*Very Large Online Platforms and Search Engines*).

⁶⁴ DSA 33. cikk (1) rendelkezés.

⁶⁵ Uo., és DSA 33. cikk (4) bekezdés.

⁶⁶ Supervision of the designated very large online platforms and search engines under DSA. European Commission, 2024. december 17. <https://tinyurl.com/376xzn37>

⁶⁷ DSA 34. cikk (1) bekezdés.

⁶⁸ Uo.

doljunk csak egyes vélemények szisztematikus vagy olykor „láthatatlan” törlésére,⁶⁹ a dezinformációs kampányokra, az online politikai propagandára és azok algoritmusok kihasználásával történő térhódítására, vagy éppen egyes demográfiai körök kirekesztésére egészségügyi információk kapcsán.⁷⁰ Amennyiben egy VLOPSE azonosít egy ilyen elfogultsági pontot, abban az esetben először értékelnie, majd csökkentenie kell az ebből eredő kockázatokat.⁷¹ Utóbbira nevesített példa, hogy tesztelniük és ellenőrizniük kell a platformoknak az alkalmazott algoritmikus rendszereiket.⁷² A DSA másik, főleg a kutatók számára előnyös rendelkezése, hogy a VLOPSE-knak, bár alapvetően szűk esetkörben, de „ismertetniük kell algoritmikus rendszereik – köztük ajánlórendszereik – tervezését, logikáját, működését és tesztelését” is,⁷³ azaz a DSA – helyesen – belátja, hogy a kutatókat szorosabban kell integrálni a VLOPSE-k kialakításából és használatából fakadó veszélyek azonosítása során.⁷⁴

Az uniós szabályozás legfontosabb fejezete az algoritmikus szabályozás tekintetében ugyanakkor minden kétséget kizárólag a jelentős várakozás után végül 2024 nyarán kihirdetett mesterséges intelligenciáról szóló rendelet (*AI Act*, „AIA”).⁷⁵ Bár az algoritmusok ritkán jelennek meg szövegszerűen a AIA-ben (összesen tíz említés), több cikk is alkalmazható az algoritmikus elfogultság jelenségére; a következőkben e szegmenseket ismertetjük. Mindenek előtt viszont egy szemantikai szempontból fontos előkérdést érdemesnek tartunk megemlíteni. Az angol nyelvű rendeletszövegben ugyanis az „*algorithmic bias*” nem jelenik meg, ugyanakkor a „*bias*” több ízben is szerepel. A magyar nyelvű fordítás ez utóbbit következetesen „torzításként” rögzíti, kissé szerencsétlen módon, hiszen a torzítás az angol szövegben konzekvensen a „*distort/distortion*” szóval jelöltetik, amelyre a magyar fordítás ugyancsak a torzítást használja.⁷⁶ Ennek említését csupán azért tartjuk relevánsnak, mert bár a következőkben citált cikkekben a torzítás szerepel *ad verbum*, jelen tanulmányban továbbra is az elfogultság kifejezéssel élünk.⁷⁷

⁶⁹ Ld. a „*shadowbanning*” jelenségét.

⁷⁰ Raj M. Ratwani – Karey Sutton – Jessica E. Galarraga: Addressing AI Algorithmic Bias in Health Care. *JAMA*, Vol. 332., No. 13. (2024). <https://doi.org/10.1001/jama.2024.13486>

⁷¹ DSA 34–35. cikkek.

⁷² DSA 35. cikk (1) bekezdés d) pont.

⁷³ DSA 40. cikk.

⁷⁴ Anna Liesenfeld: The Legal Significance of Independent Research based on Article 40 DSA for the Management of Systemic Risks in the Digital Services Act. *European Journal of Risk Regulation*, (2024) 6–8. <https://doi.org/10.1017/err.2024.61>

⁷⁵ Az Európai Parlament és a Tanács (EU) 2024/1689 rendelete (2024. június 13.) a mesterséges intelligenciára vonatkozó harmonizált szabályok megállapításáról, valamint a 300/2008/EK, a 167/2013/EU, a 168/2013/EU, az (EU) 2018/858, az (EU) 2018/1139 és az (EU) 2019/2144 rendelet, továbbá a 2014/90/EU, az (EU) 2016/797 és az (EU) 2020/1828 irányelv módosításáról (a mesterséges intelligenciáról szóló rendelet) (EGT-vonatkozású szöveg) PE/24/2024/REV/1 HL L, 2024/1689, 12.7.2024.

⁷⁶ Ehhez ld. AIA (29) preambulumbekkezdés.

⁷⁷ Többek között azért is ragaszkodunk ehhez, mert a magyar fordítás számos szemantikailag nehezen értelmezhető terminust használ, így például nehezen interpretálható a „méltánytalan torzítás” (*unfair bias*, (27) preambulumbekkezdés), amely jelentéstartalmát tekintve jóval inkább az olyan túlzó módon elfogult algoritmikus folyamatokat takarja, amelyek valamilyen módon negatív hatással járnak a fel-

Az elfogultság elsőként az AIA 27. preambulumbekzdésében jelenik meg, amely kiemeli, hogy az algoritmikus elfogultságot a sokszínűség, a méltányosság és a megkülönböztetésmentesség előmozdításával kell kezelni a mesterséges intelligencia rendszerekben. A hivatkozott preambulumbekzdés utal a 2019. április 8-án a mesterséges intelligenciával foglalkozó magas szintű uniós szakértői csoport által bemutatott megbízható mesterséges intelligenciára vonatkozó etikai iránymutatásokra is, amelyek a befogadó fejlődést és a tisztességtelen vagy diszkriminatív hatások elkerülését szorgalmazzák.⁷⁸ Az iránymutatások, ellentétben az AIA-val, lefektetik az elfogultság fogalmát, amely szerint a jelenség olyan tendenciákra vagy előítéletekre vonatkozik, amelyek (1) befolyásolhatják az kimeneteleket (*outputs*), és (2) amelyek különböző forrásokból, például az adatgyűjtésből, a szabálytervezésből, a felhasználói interakcióból vagy korlátozott alkalmazási kontextusból erednek.⁷⁹ Az iránymutatás – helyesen – azt is aláhúzza, hogy míg az elfogultság „természetét” tekintve lehet szándékos vagy nem szándékos, és bizonyos esetekben akár előnyös is, az esetek többségében az algoritmusok torzító jellege diszkriminatív eredményekhez vezethet – ezt az iránymutatás tisztességtelen elfogultságnak (*unfair bias*) nevezi.⁸⁰ A tisztességtelen elfogultság elkerülése érdekében az iránymutatás hangsúlyozza az adekvát problémakezelés fontosságát, így például az adatokban meglévő (inherens) előítéletek és a diszkriminatív algoritmikus-tervezés elkerülését és orvoslását.⁸¹ Ugyan az iránymutatás erősen absztraháltan fogalmaz, az „enyhítési stratégiák” közé sorolja az átlátható felügyelet biztosítását, valamint a fejlesztőcsapatok sokszínűségének előmozdítását.⁸²

Az AIA 48. preambulumbekzdése aláhúzza, hogy a mesterséges intelligencia rendszerek (MI-rendszerek⁸³) potenciálisan káros hatással lehetnek az alapvető jogokra, beleértve a megkülönböztetésmentességet, az egyenlőséget és a méltányosságot.⁸⁴ Bár az AIA kockázatalapú-struktúrájának részletes ismertetése meghaladná e dolgozat kereteit, aláhúzendó, hogy – akárcsak a DSA esetében – az MI-rendszereknél is differenciált rendszert alkalmaz a rendelet. Ennek értelmében az MI-rendszereket azok kockázata szerint osztályozza a jogszabály, ahol a tiltott, a nagy kockázatú és az általános, de rendszerszintű kockázatot jelentő MI-rendszerekre más szabályok vonatkoznak. Ugyan a kockázat-rendszer, amely a rendeletjavaslatban distinkatív módon, négy részre volt különítve, jelentősen változott a kihirdetett szövegben,⁸⁵ az algoritmusok szerepe

használóra nézve (beleértve a méltányosság kérdését is). Azaz, tulajdonképpen az „előítéletességre”, mintsem a torzításra utal.

⁷⁸ A teljes dokumentum elérhető itt: European Commission: Ethics guidelines for trustworthy AI. 2019. április 8. <https://tinyurl.com/2nyzyy6k>

⁷⁹ Uo. 38.

⁸⁰ Uo.

⁸¹ Uo. 18.

⁸² Uo.

⁸³ AIA 3. cikk 1. pont., vö. (28) prb.

⁸⁴ Vö. A Biden-féle kormányzat kapcsán korábban említett *executive order*.

⁸⁵ Claudio Novelli – Federico Casolari – Antonino Rotolo – Mariarosaria Taddeo – Luciano Floridi: Taking AI risks seriously: a new assessment model for the AI Act. *AI & Society*, Vol. 39., No. 5. (2023) 2493–2497. <https://doi.org/10.1007/s00146-023-01723-z>

és jelentősége a jogszabályban összességében – tartalmát tekintve – megegyezik a korábbi szövegváltozatokkal. E kontextusban a tiltott MI-rendszerek tekintetében az EU területén tilosak az olyan MI-gyakorlatok, így az olyan algoritmusok implementálása, amelyek súlyosan beavatkoznak és torzítják az emberi döntéshozatalt, kihasználják a felhasználó „sebezhetőségét”, osztályozzák az embereket bizonyos kritériumok szerint (ez a passzus egyértelműen a kínai kreditrendszerre utal), bűnelkövetés megijósításához és arcfelismerő adatbázisok létrehozásához használhatóak, természetes személyek érzelmeiből következtetéseket vonnak le,⁸⁶ biometrikus kategorizálást alkalmaznak, és „valós idejű” távoli biometrikus azonosító rendszereket használva a nyilvánosság számára hozzáférhető helyeken bűnüldözési célokat szolgálnak.⁸⁷ A biometrikus azonosítás tekintetében külön ki kell emelni a 32. preambulumbekendést, amely tételesen utal az elfogultság kérdésre.⁸⁸ A bekezdés értelmében a valós idejű biometrikus azonosítás által létrejövő „torzított eredmények és diszkriminatív hatások különösen relevánsak az életkor, az etnikai és a faji hovatartozás, a nem vagy a fogyatékoságok tekintetében,”⁸⁹ így ezért is tiltott (*lex generalis*) az alkalmazása.⁹⁰

A nagy kockázatú MI-rendszereknél, a DSA-hoz hasonló módon, az AIA is bevezet egy kifejezetten az érintett MI-gyakorlat kialakítására és működésére vonatkozó kockázatértékelési rendszert, amely az adott rendszer teljes életciklusa során megköveteli a kockázatok folyamatos azonosítását, elemzését és mérséklését, beleértve az algoritmusok elfogultságát is.⁹¹ Bár *expressis verbis* az algoritmusokat nem említi, tartalmát tekintve az AIA 13. cikke, amely a nagy kockázatú rendszerek átláthatóságáról rendelkezik, nagy mértékben értelmezhető az algoritmusok transzparenciájára vonatkozó szabályként is. E körben az AIA megköveteli, hogy legyen világos és átfogó dokumentáció az érintett mesterséges intelligencia rendszer képességeinek és korlátainak magyarázatára, beleértve annak lehetőségét, hogy elfogult eredményeket produkáljon. Érdemes kiemelni e körben a (67) preambulumbekendésben részleteiben tárgyalt „tanításhoz, validáláshoz és teszteléshez használt adatkészletek” kérdését. E körben az algoritmusok elfogultság kérdése az olyan adattömegek kapcsán merül fel, amelyek kifejezetten személyekre vagy csoportokra vonatkozó statisztikai adatokat tartalmaznak és amelyek ezáltal potenciálisan kirekesztőek lehetnek „bizonyos kiszolgáltatott csoportokkal” szemben. A kockázatok mérsékléséhez az ilyen adatkészleteknek rendelkeznie kell azokkal a jellemzőkkel és tulajdonságokkal, amelyek „azon sajátos földrajzi, kontextuális, magatartási vagy funkcionális környezethez kapcsolódnak, amelyben az MI-rendszert használni szándékozzák.” A vonatkozó 10. cikk pedig konkrétan lefek-

⁸⁶ Ehhez bővebben ld. AIA 44. preambulumbekendés, amely külön kihangsúlyozza az általánosíthatóság problémáját.

⁸⁷ AIA 5. cikk (1) bekezdés. Fontos ugyanakkor megemlíteni, hogy egyes tiltott MI-gyakorlatok esetében, így például a bűnüldözési célokra használható rendszerek tekintetében fennállhatnak kivételes esetek, amikor e rendszerek használata megengedett.

⁸⁸ „[...] természetes személyek távoli biometrikus azonosítására szolgáló MI-rendszerek technikai pontatlansága torzított eredményekhez vezethet, és diszkriminatív hatásokat eredményezhet”

⁸⁹ Uo.

⁹⁰ Ugyancsak ld. AIA (54) prb.

⁹¹ AIA 9. cikk.

teti, hogy az ilyen nagy kockázatú MI-rendszereknek (magyarán olyanoknak, amelyek a tanító, a validálási és a tesztadatkészletekkel dolgoznak), kötelezően szükséges minden olyan lehetséges kockázat kivizsgálása, amelyek az alapvető jogokat sérthetik, megkülönböztetéshez vezetnek, illetve ha bármilyen tekintetben hatnak (azaz nemcsak negatív, hanem akár semleges vagy pozitívan is) az érintett személyek egészségére és biztonságára.⁹² E körben hangsúlyos a kivizsgálás akkor, ha „az adatok kimenetei befolyásolják a jövőbeli műveletek bemeneteit,” azaz például olyan esetekben ahol korábbi adatokra építve tesz majd az MI később javaslatokat; hipotetikusan például olyan városrészekben javasol magasabb rendőri jelenlétet, ahol korábban több bűnözés volt. A kivizsgálás mellett az ilyen MI-rendszereknél szükséges preventív és enyhítő intézkedéseket is bevezetni.⁹³ Ezen túlmenően, a 10. cikk 5. pontja kiemeli, hogy a magas kockázatot jelentő MI-rendszerek szolgáltatói a személyes adatok különleges kategóriáit (pl. faji, egészségügyi vagy politikai meggyőződésre utaló adatokat) csak és kizárólag akkor kezelhetik, ha az elfogultságok felderítéséhez és korrigálásához feltétlenül szükségesek továbbá, ha az elfogultságok felderítése semmilyen más módon nem lehetséges.⁹⁴ Az ilyen adatfeldolgozásnak szigorú biztosítékokat kell követnie, beleértve az álnevesítést, a korlátozott hozzáférést, a biztonsági ellenőrzéseket, a harmadik felekkel való megosztás tilalma és az időben történő törlés.⁹⁵ Továbbá részletes dokumentációval kell igazolni azt is, hogy az érzékeny adatok felhasználása miért volt szükséges, illetve bizonyítani kell az uniós adatvédelmi jogszabályoknak való megfelelést. Kiemelendők még a magas kockázatú MI-rendszereknél a 14. cikk rendelkezései, amelyek aláhúzzák, hogy az ilyen rendszereknek hatékony emberi felügyeletet kell biztosítaniuk az egészséget, a biztonságot és az alapvető jogokat érintő kockázatok minimalizálása érdekében. Az algoritmikus elfogultság kontextusában kulcsfontosságú a 4. pont b), amely kifejezetten az automatizálási torzítás kérdését érinti.⁹⁶ E körben a rendelet megköveteli, hogy a felügyeleti mechanizmusok segítsék a felhasználókat abban, hogy tudatában legyenek ennek a kockázatnak, lényegében annak biztosítékaként, hogy az emberek képesek legyenek megkérdőjelezni vagy akár felülbírálni a mesterséges intelligencia döntéseit. A tervezés és az elfogultság kapcsán a 15. cikk szolgáltáspontként. E cikk kiemeli, hogy a nagy kockázatú MI rendszereket úgy kell megtervezni, hogy minimalizálják az elfogult visszacsatolási hurkokat,⁹⁷ különösen azokban a rendszerekben, amelyek a telepítés után is folytatják a tanulást. Ez magában foglalja olyan technikai és szervezeti intézkedések végrehajtását is, amelyek megaka-

⁹² AIA 10. cikk 2. pont f) bekezdés.

⁹³ AIAI 10. cikk 2. pont g) bekezdés.

⁹⁴ Részletes elemzéshez ld. Marvin van Bekkum: Using sensitive data to de-bias AI systems: Article 10(5) of the EU AI act. *Computer Law & Security Review*, Vol. 56. (2025) <https://doi.org/10.1016/j.clsr.2025.106115>

⁹⁵ Ld. Uo. AIA 5.pont (a)–(f).

⁹⁶ Ez egy tendenciát jelent, amelynek lényege, hogy az emberek túlzottan is az MI által létrehozott eredményekre támaszkodnak.

⁹⁷ Általánosságban véve azt értjük rajta, amikor egy MI-rendszer kimenetei a jövőbeli tanulás vagy döntéshozatal bemeneteként szolgálnak.

dályozzák, hogy az elfogult kimenetek negatívan befolyásolják a jövőbeli bemeneteket, és biztosítják az ilyen kockázatok aktív csökkentését is.

5. Az algoritmikus elfogultság mint lernéi Hüdra és a jövő szabályozási kérdései

Mitológiai hasonlattal élve, az MI, különösen az algoritmusok szabályozása, hasonlít Héraklész párbajára a lernéi Hüdrral – minden szabályozási próbálkozásra két újabb, addig ismeretlen vagy jogi eszközökkel kezelhetetlen probléma kerül elő. Az egyik legfontosabb, továbbra is megoldásra váró polémia az algoritmusok és adatkészletek kialakításában rejlő elfogultság. Kétségtelen, hogy akár a DSA, akár az AIA által bevezetett átláthatósági szabályok fontos lépések, ugyanakkor továbbra is fennáll az algoritmusok kialakításakor megjelenő feketedoboz („*black box*”)-hatás, amely tulajdonképpen az MI-rendszerekbe inherensen „kódolt” átláthatatlanságot jelenti.⁹⁸ A feketedobozok átláthatatlansága lehetetlenné teszi a probléma gyökerének megtalálását, azaz a elfogultságok azonosítását és orvoslását, még akkor is, ha a rendeletek által előírt dokumentációs kötelezettség teljesülne.

További akadály lehet a végrehajtás. Bár különösen az AIA lefektet végrehajtási szabályokat, kérdéses, hogy biztosítható-e megfelelő forrással és leginkább szakértelemmel a jogszabály megfelelő betartatása. E tekintetben kifejezetten aggasztó az MI-hivatal, amelynek célja, hogy hozzájáruljon „az MI-rendszerek és az általános célú MI-rendszerek, valamint az MI-irányítás végrehajtásához, nyomon követéséhez és felügyeletéhez”.⁹⁹ Az elmúlt egy évben elsősorban adminisztratív eredményeket tud felmutatni, míg a Hivatal munkáját támogató MI-testület eddig összesen kétszer ülésezett, érdemi eredmény nélkül.¹⁰⁰ Az erőforrások hiánya és a lassú bürokratikus folyamatok még inkább láthatóak nemzeti szinten, ugyanis a tagországok többségében egyelőre nincs kifejezetten az MI-re és annak ellenőrzésére szakosodott hatóság. Az erőforrás és szakértelem e körben messzemenőleg túlmutat az AIA ismeretén; súlyos erőforráshiány van mind technológiai, mind akár MI-etikai szakértőből.¹⁰¹

Talán még nagyobb, és a jogi szabályozáson túlmutató probléma az algoritmusok szabályozásának összehangolatlansága. Ahogy az ENSZ Főtitkára által összeállított MI-vel foglalkozó szakértői panel is jelezte 2024 őszén: „megdönthetetlen” szükséglet egy globális szabályozás kialakítása és nem szabad engedni, hogy kizárólag a piaci erők diktálják az MI-fejlesztést és a szabályozás korlátait.¹⁰² A több, mint száz oldalas riport arra az aggodalomra is kitér, hogy megfelelő, globális kormányzás nélkül a mes-

⁹⁸ Bartosz Brożek – Michał Furman – Marek Jakubiec – Bartłomiej Kucharzyk: The black box problem revisited. Real and imaginary challenges for automated legal decision making. *Artificial Intelligence and Law*, Vol. 32., No. 2. (2023) 427–428. <https://doi.org/10.1007/s10506-023-09356-9>

⁹⁹ AIA 3. cikk 47. pont.

¹⁰⁰ Ld. European AI Office. *EU*, 2024. <https://digital-strategy.ec.europa.eu/en/policies/ai-office>

¹⁰¹ Az egyik legnagyobb számítástechnikai újság, a *ComputerWeekly* felmérése szerint az európai IT-ban dolgozók 40%-ának jelentős hiányosságai vannak az AI kapcsán. Clare McDonald: AI and cyber skills worryingly lacking, say business leaders. *ComputerWeekly*, 2024. július 11. <https://tinyurl.com/27kn7p96>

¹⁰² Governing AI for Humanity. *UN Final Report*, 2024. szeptember. 7–8.

terséges intelligencia előnyei csak néhány kiválasztott nemzetre és szervezetre korlátozódhatnak, amely tovább súlyosbíthatja az eddig is fennálló digitális megosztottságot és az egyenlőtlenségeket.¹⁰³ A harmonizáltság idealisztikus, vagy talán romanticizált képe viszont a gyakorlatban igen képlékeny talajon áll. Sajnálatos módon ugyanis a fentebb említett anyag nem tér ki arra, hogy az algoritmikus elfogultság nem csak és kizárólag technológiai kérdés; sőt, számos olyan társadalmi és kulturális elemet is hordoz magában (akár közvetve vagy közvetlenül), amelyek a globális szabályozással nem tűnnek feloldhatónak. A szabályozási harmonizáció ugyanis jelen esetben egy globális „társadalmi elv” elfogadását is jelenti, amely az emberi jogok és a méltányosság értelmezése tekintetében is közös értékeket feltételez – helytelenül és alaptalanul. Kritikai szemszögből vizsgálva a harmonizáció kérdését megállapítható az ENSZ-jelentés kapcsán, hogy a javasolt normakezdeményezések a nyugati, demokratikus elvekből erednek, amelyek nem feltétlenül elfogadottak vagy érvényesülnek általánosan – gondolhatunk itt olyan régiók országaira, ahol a nők és marginalizált közösségek védelme folyamatosan sérül. Így a diszkrimináció csökkentésére törekvő nemzetközi normák végérvényesen egy általános jogalkotói és társadalmi attitűdöt is közvetítenek, amelyek bizonyos szocio-kulturális kontextusokban ütközhetnek az uralkodó intézményi vagy ideológiai struktúrákkal. Végül érdemes a jövőben azt is figyelembe venni, különösen a fentiek fényében, hogy az algoritmikus elfogultság globalizált szabályozása egyáltalán szolgálhat-e, sőt, kellene-e szolgálnia az etikai értékek technikai szabályozásnak álcázott exportjának eszközeként. Ennek vizsgálatához fontos lehet a normatív imperializmus elkerülése,¹⁰⁴ amely a helyi társadalmi-politikai realitásoktól függetlenül a nyugati prioritásokat beágyazza a határokon átnyúlóan működő algoritmusokba.

6. Következtetések

A fenti metaforát folytatva, Héraklésznak szüksége volt Iolaoszra, aki (bár némi isteni segítséggel) megtalálta a megfelelő fogást a Hüdrán, s így segítségével a thébai hős legyőzte a szörnyet. Az algoritmikus elfogultság szabályozásánál, egyetértve az ENSZ riportjával, Iolaosz nem egy jogszabály, egy újszerű rendelet vagy akár az MI-fejlesztők saját szabályozása, hanem jóval inkább egy globális szintű, átfogó, kooperatív szabályozás lehet. Ez a kooperatív munka sokkal többet takar, mint egyes államok közös erőfeszítését; be kell vonni a civil szervezeteket, a kutatókat, a különböző érdekképviseleti szervezeteket, az MI-ben aktívan részt vevő gazdasági szereplőket és természetesen hangot kell adni a marginalizált csoportoknak is, akiket a leginkább érint az algoritmusok kódolt kirekesztése.

Jelen tanulmányban az algoritmikus elfogultság kérdését vizsgáltuk, különös tekintettel annak fogalmi alapjaira és szabályozás jelenlegi helyzetére. A dolgozatban ismertettük a legfontosabb jogi irányokat, és kiemelt hangsúlyt fektettünk az európai (uniós) szabályok bemutatására. A kutatás kiemelt eredménye – egyben fontos felhívása –,

¹⁰³ Uo. 27.

¹⁰⁴ Ld. Julian Pänke: The Fallout of the EU's Normative Imperialism in the Eastern Neighborhood. *Problems of Post-Communism*, Vol. 62., No. 6. (2015) 350–363. <https://doi.org/10.1080/10758216.2015.1093773>

hogy sürgősen szükség van egy harmonizált, globális szabályozási keretre. A jelenlegi regionális és nemzeti erőfeszítések ugyanis – bár dicséretesek – nem elegendőek a mesterséges intelligencia fejlesztésének és alkalmazásának határokon átnyúló jellegének kezelésére. A jövőbeni kutatásokra javasolt az algoritmikus elfogultsággal kapcsolatos technikai megoldások bemutatása, a jogi kérdések tekintetében pedig fontos eredményekkel szolgálhatnak majd a hazai irányokat előrevetítő vagy akár az EU-n és Egyesült Államokon kívüli szabályozási kezdeményezések és megoldások bemutatása.

ALGORITMUSOK ÉS ÁTLÁTHATÓSÁG

Az online platformok ajánlórendszereinek jogi szabályozása az Európai Unióban

Mezei Kitti*

1. Bevezetés: Ajánlott tartalom vagy irányított figyelem? Az Európai Unió új szabályozási megközelítése

A digitális szolgáltatások térnyerése alapjaiban alakította át az online információ-áramlást, a tartalomfogyasztási szokásokat és a platformok működését. Az Európai Unió által elfogadott digitális szolgáltatásokról szóló rendelet (a továbbiakban: DSA)¹ célja, hogy szabályozza az online térben működő szolgáltatások jogi környezetét, biztosítva a felhasználók jogainak védelmét, a transzparenciát és a digitális piac kiegyensúlyozott működését.

Az online ajánlórendszerek különös figyelmet kaptak a szabályozásban, mivel alapvető szerepet játszanak a digitális tartalom elérhetőségének és láthatóságának meghatározásában. Az algoritmikus döntéshozatal² révén az ilyen rendszerek befolyásolják a felhasználók által látott információkat, így közvetetten hatással

* Tudományos főmunkatárs, HUN-REN Társadalomtudományi Kutatóközpont, Jogtudományi Intézet. ORCID: <https://orcid.org/0000-0002-6497-1186> A kutatás a Magyar Tudományos Akadémia Bolyai János Kutatási Ösztöndíjának támogatásával valósult meg.

¹ Az Európai Parlament és a Tanács (EU) 2022/2065 rendelete (2022. október 19.) a digitális szolgáltatások egységes piacáról és a 2000/31/EK irányelv módosításáról (digitális szolgáltatásokról szóló rendelet), HL L 277., 2022.10.27. pp. 1–102.

² Az „algoritmus” kifejezés nem értelmezhető akként, mintha az algoritmus önálló döntéshozatali képességgel rendelkezne. A jogalkotó csupán egy általánosan elterjedt technikai terminológiát vett át. Az algoritmikus döntéshozatal alatt minden olyan automatizált mechanizmust érteni kell, amely során a döntéshozatal legalább részben automatizált módon, emberi beavatkozás nélkül történik. Ennek megfelelően nem feltétele a fogalom alkalmazhatóságának, hogy az érintett eljárás különösen komplex algoritmikus struktúrán alapuljon, vagy megfeleljen az Európai Bizottság által az AI Act-ben meghatározott mesterséges intelligencia rendszerekre vonatkozó feltételeknek. Már az is elég, ha valamilyen egyszerű, automatizált megoldást alkalmaznak, például egy egyszerű szabályalapú programot. Ld.



vannak a közvélemény alakulására, a piac versenyhelyzetére és az alapvető jogok érvényesülésére. Az EU jogalkotása kiemelten kezeli az ajánlórendszerek átláthatóságát és elszámoltathatóságát, különös tekintettel a rangsorolási mechanizmusokra, a célzott tartalomajánlásokra és az automatizált döntéshozatali eljárásokra.

A tanulmány célja, hogy részletesen bemutassa a DSA ajánlórendszerekkel kapcsolatos rendelkezéseit, elemezze azok gyakorlati következményeit, valamint jogi és etikai szempontból értékelje az új szabályozási keret hatásait. Az elemzés során különös figyelmet fordítok a platformszolgáltatók kötelezettségeire, az átláthatósági követelményekre, valamint azokra a jogorvoslati mechanizmusokra, amelyek biztosítják a felhasználók jogainak védelmét a digitális térben.

2. Az ajánlórendszerek egyes etikai és társadalmi kihívásai

Az egyik sokat vitatott etikai kérdés, az ajánlórendszerekkel kapcsolatban az egyéni döntési folyamatokba való beavatkozás mértéke és annak hatása. Az ajánlórendszerek nem közvetlenül elrejtik, hanem kevésbé láthatóvá teszik a nem annyira relevánsnak ítélt lehetőségeket, tartalmakat, például azáltal, hogy azokat nem helyezik az első találatok közé. Ennek eredményeként a felhasználók nagyobb valószínűséggel fogadják el az első ajánlásokat, még akkor is, ha azok nem feltétlenül felelnek meg az egyéni preferenciáiknak.³

Ez a „*position bias*”, amelynek hatása kettős. Egyrészt manipulációs kockázatok merülnek fel, ha az ajánlások pontatlanok vagy üzleti érdekek által manipuláltak, a felhasználók nem a saját érdekeiknek megfelelő döntéseket hozzák meg.⁴ Másrészt hiányos információkhoz vezethet, ha a felhasználók csak egy szűk keresztmetszet alapján formálnak véleményt, így fontos információk elkerülhetnek a figyelmüket. Ez felveti a kérdést, hogy az első találatok között szándékosan meg kellene-e jeleníteni kevésbé relevánsnak ítélt lehetőségeket – ezt a gyakorlatot „*serendipity by design*” néven említik.⁵ További etikai problémát jelent a beépített ösztönző technikák (*nudging*) alkalmazása. Ezek például sürgethetik a felhasználókat, hogy gyorsan döntsenek. Ugyanakkor fontos kiemelni, hogy ezek az algoritmusok és technikák nemcsak kockázatokat, hanem előnyöket is hordoznak. A platformok hosszú távú érdeke, hogy valóban releváns és hasznos ajánlásokat tegyenek, hiszen a felhasználói bizalom és hűség ezen múlik. A jól

erről Benjamin Raue: Art 14 – Allgemeine Geschäftsbedingungen. In: Franz Hofmann – Benjamin Raue (szerk.): *Digital Services Act*. Nomos, Baden-Baden, 2023. 259.

³ Sergio Genovesi – Katharina Kaesling – Scott Robbins: Introduction: Understanding and Regulating AI-Powered Recommender Systems. In: Sergio Genovesi – Katharina Kaesling – Scott Robbins (szerk.): *Recommender Systems: Legal and Ethical Issues*. (The International Library of Ethics, Law and Technology 40) Cham, Springer, 2023. 2. https://doi.org/10.1007/978-3-031-34804-4_1

⁴ Ld. erről Fleur Jongepier – Michael Klenk (szerk.): *The Philosophy of Online Manipulation*. (Routledge Research in Applied Ethics) Routledge, 2022.; valamint Daniel Susser – Beate Roessler – Helen Nissenbaum: Technology, Autonomy, and Manipulation. *Internet Policy Review*, Vol. 8., No. 2. (2019).

⁵ Urbano Reviglio: Serendipity by Design? How to Turn from Diversity Exposure to Diversity Experience to Face Filter Bubbles in Social Media. In: *International Conference on Internet Science*. Cham, Springer, 2017. 281–300. https://doi.org/10.1007/978-3-319-70284-1_22

működő ajánlórendszerek segíthetnek az információúterhelés kezelésében, és hatékonyabbá tehetik a döntéshozatalt azáltal, hogy a felhasználó számára legvalószínűbben releváns lehetőségeket emelik ki. Az ösztönző technikák pedig nem feltétlenül manipulációs szándékúak, megfelelő alkalmazásuk esetén valóban támogatják a felhasználót a döntéseiben.⁶

Az ajánlórendszerek alapvetően két típusba sorolhatók, amennyiben a szűrési módszerek képezik a megkülönböztetési kritériumot: tartalom alapú és együttműködésen alapuló szűrés. Ezeket kiegészíti egy harmadik kategória, a hibrid módszerek.⁷

A tartalom alapú szűrés olyan tartalmakat ajánl, amelyek hasonlóságot mutatnak a felhasználó korábbi preferenciáival.⁸ Ez a módszer különösen videó- és streaming platformokon alkalmazható, ahol általában két algoritmus működik, az egyik egy lehetséges ajánlási „poolt” generál, míg a másik értékeli és rangsorolja az egyes tartalmakat ebből a körből. Az együttműködésen alapuló szűrés olyan tartalmakat ajánl, amelyek más, hasonló értékelési szokásokkal rendelkező felhasználókat is érdekelnek.⁹ Ebben az esetben nincs szükség további információra az ajánlott tartalomról; az algoritmus a felhasználók viselkedése alapján működik. Ezt a módszert jellemzően kereskedelmi platformokon alkalmazzák, ahol a rendszer implicit visszajelzésekre (például tranzakciók vagy böngészési előzmények) vagy explicit interakciókra (például korábbi értékelések) épít. A hibrid szűrési módszerek kombinálják az előbbieket: először külön alkalmazzák az együttműködésen és tartalom alapú szűrési eljárásokat, majd ezek alapján előrejelzéseket készítenek a felhasználói viselkedésről. Az ilyen hibrid ajánlórendszerek különösen közösségi média platformokon találhatók meg, ahol az ajánlások nemcsak tartalmi alapúak, hanem a felhasználók társadalmi viselkedését is figyelembe veszik.¹⁰

A közösségi média platformokon az egyik legerősebb befolyásoló tényező a tartalom-ajánló rendszer, amely meghatározza a felhasználók számára megjelenített tartalmak sorrendjét. Ez egyfajta kapuőr szerepet tölt be, mivel a tartalomajánló rendszer – például a Facebook, Instagram vagy TikTok algoritmusai – határozzák meg, hogy a felhasználók milyen bejegyzésekkel találkoznak a hírfolyamukban, és melyeket nem is látnak egyáltalán. Ez a válogatás jelentős befolyással van arra, hogy milyen információk jutnak el az emberekhez, ezért egyre több közéleti vita középpontjába kerül.¹¹ Az ajánlások

⁶ Ld. bővebben: Francesco Ricci – Lior Rokach – Bracha Shapira (szerk.): *Recommender Systems Handbook* (2nd ed.) Cham, Springer, 2015.

⁷ Dagmar Gesmann-Nuissl – Sebastian Meyer: Black Hole Instead of Black Box? The Double Opaqueness of Recommender Systems on Gaming Platforms and Its Legal Implications. In: Genovesi–Kaesling–Robbins (szerk.) i. m. 64–65. https://doi.org/10.1007/978-3-031-34804-4_4

⁸ Gediminas Adomavicius – Alexander Tuzhilin: Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions. *IEEE Transactions on Knowledge and Data Engineering*, Vol. 17., No. 6. (2005) 734–749. <https://doi.org/10.1109/TKDE.2005.99>

⁹ Uo.

¹⁰ Yang Wang – Michal Kosinski – David Stillwell – John Rust – Thore Graepel: Modeling and Predicting Personality with Social Media. *International Conference on Artificial Intelligence and Statistics (AISTATS)*, Vol. 31. (2013).

¹¹ Jennifer Cobbe – Jatinder Singh: Regulating Recommender Systems: Transparency, Fairness, and Accountability. *Computer Law & Security Review*, Vol. 35., No. 5. (2019) 105342. <https://doi.org/10.1016/j.clsr.2019.105342>

áltában a kezdőoldalon jelennek meg, például Facebook esetén beágyazva a barátok bejegyzései közé, de a hírfolyam sorrendjét rangsorolási algoritmusok határozzák meg.¹²

Két alapvetően eltérő ajánlási elv figyelhető meg: érdeklődésalapú ajánlások – ezek célja a felhasználó és a tartalom közötti relevancia értékelése, hogy a leginkább érdeklődésre számot tartó tartalom jelenjen meg (ez a tartalom alapú ajánlás középpontja). Ez a modell a felhasználó korábbi interakcióira – például kedvelésekre, keresésekre, megtekintési időre – és a tartalmak jellemzőire (kulcsszavak, témák, metaadatok) épít. A cél annak az értékelése, hogy az adott tartalom mennyire releváns az adott felhasználó számára. A YouTube főoldala személyre szabott videókat kínál, a felhasználó által korábban megtekintett tartalmak témája, hossza, stílusa alapján. Az Instagram Explore algoritmus az érdeklődési körök alapján (például hashtagek, profilok, keresési előzmények) generál vizuálisan releváns javaslatokat.

A befolyásalapú ajánlások esetén azt vizsgálják, hogy milyen tartalmakat osztanak meg a lehető legnagyobb hatás elérése érdekében (ez az együttműködésen alapuló szűréshez kapcsolódik).¹³ A befolyásalapú logika az egyes tartalmak kollektív népszerűségére és a tömeges megosztások és széles körű láthatóság mechanizmusaira épít. A rendszer más felhasználók – különösen azok, akik hasonló érdeklődéssel vagy kapcsolati hálóval rendelkeznek – viselkedésmintáiból következtet a javasolható tartalmakra. A Facebook hírfolyam-rangsorolása figyelembe veszi, hogy az ismerősök közül kik reagáltak egy adott posztra – így egy tartalom előtérbe kerülhet a kollektív aktivitás révén, nem pedig személyes preferencia alapján is. A TikTok esetében a tartalom széles körben terjedése (megosztások száma, kommentek, gyors terjedés) szintén befolyásolja annak algoritmikus felerősítését – még olyan felhasználók számára is, akik korábban nem érdeklődtek hasonló tartalom iránt.

A mesterséges intelligencia alapú rendszer¹⁴ átláthatatlansága (*opacity*) további komoly etikai problémát jelent számos algoritmus esetén, így nincs ez másként az ajánlórendszerek esetében sem. Ez akkor fordul elő, ha a döntéshozatali folyamat nem átlátható, és a felhasználók számára nincs lehetőség annak megértésére, hogy milyen tényezők befolyásolják az ajánlásokat. Például, ha egy szálláskereső rendszer olyan információkat kér be, amelyek nem kapcsolódnak egyértelműen a keresés pontosításához, de nem magyarázza el azok feldolgozását, ez végső soron akár átláthatósági és adatvédelmi problémát is jelenthet a felhasználó oldaláról nézve.¹⁵

Az átláthatatlan döntések megakadályozzák a felhasználókat abban, hogy átlássák és megértsék azokat, ezáltal csökken az emberi kontroll és az autonómia. Ezzel szemben az átlátható rendszerek megerősítik a felhasználói jogokat, lehetővé téve a döntések

¹² Paddy Leerssen: The Soap Box as a Black Box: Regulating Transparency in Social Media Recommender Systems. *European Journal of Law and Technology*, Vol. 11., No. 2. (2020).

¹³ Gesmann-Nuissl–Meyer i. m. 65–66.; valamint Wang et al. i. m.

¹⁴ AI Act 3. cikk 1. pont értelmében a mesterséges intelligencia rendszer olyan „gépi alapú rendszer, amelyet különböző autonómiaszinteken történő működésre terveztek, és amely a bevezetését követően alkalmazkodóképességet tanúsíthat, és amely a kapott bemenetből – explicit vagy implicit célok érdekében – kikövetkezteti, miként generáljon olyan kimeneteket, mint például előrejelzéseket, tartalmakat, ajánlásokat vagy döntéseket, amelyek befolyásolhatják a fizikai vagy a virtuális környezetet.”

¹⁵ Genovesi–Kaesling–Robbins i. m. 2–4.

felülvizsgálatát. Az informatika területén aktív vita zajlik arról, hogy a gépi tanuláson alapuló automatizált döntéseket milyen mértékben kell és lehet egyáltalán megmagyarázni.¹⁶ Ez különösen akkor válik problémássá, amikor a felhasználó számára nem világos, hogy milyen tényezők befolyásolják a kapott ajánlásokat, és nincs lehetősége azok háttérének megértésére vagy kontrollálására. Vegyük például a szálláskereső platformokat, mint a Booking.com vagy az Airbnb. Ezek a rendszerek gyakran kérnek be olyan személyes vagy technikai adatokat – például a felhasználó tartózkodási helyét, eszközeinek típusát, vagy korábbi keresési szokásait –, amelyeknek a felhasználó szempontjából nincs egyértelmű kapcsolata a keresés finomhangolásához. Előfordulhat, hogy egy Apple-terméket használó felhasználónak eleve drágább szállásopciókat mutat előbb, mint egy másik eszközzel keresőnek. A felhasználó számára viszont nem derül ki, hogy az eszköz típusa hogyan befolyásolta az ajánlásokat vagy az árakat, így elveszik a döntéshozatal tudatossága.¹⁷ Hasonló problémák jelennek meg a tartalomajánló rendszerek esetén is, például a Netflix vagy a YouTube esetében. Egy felhasználó számára nem világos, hogy egy adott filmajánlás mögött a korábbi megtekintések, más felhasználók viselkedése, vagy esetleg üzleti megfontolások (például egy adott tartalom kiemelt promóciója) állnak. Emiatt sokan kérdőjelezik meg az ajánlások relevanciáját, és gyakran érzik úgy, hogy nem valós érdeklődésükre, hanem rejtett üzleti érdekekre alapulnak a javaslatok.

Érdekesként megemlítenéd, hogy a videójátékplatformokkal kapcsolatban a recens szakirodalom¹⁸ „fekete doboz hatás” helyett „fekete lyukról” beszél, mivel ezek multidimenzionális rendszerekként funkcionálnak. Például a Steam interaktív ajánló-rendszere nem csupán a címkék és értékelések alapján működik, hanem folyamatosan tanul a felhasználók játékszokásaiból, összevetve azokat más játékosok viselkedésével. Ennek eredményeként a felhasználók számára személyre szabottabb ajánlások készülnek, amelyeket tovább finomíthatnak azáltal, hogy barátaik preferenciáit is figyelembe veszik. A játékrendszerek alapvetően hasonló módon működnek, mint más ajánló-rendszerekkel rendelkező platformok (például Amazon vagy YouTube). A különbség abban rejlik, hogy többféle algoritmus kombinációját alkalmazzák, és bizonyos kontextusokban egyes adatokat figyelmen kívül hagynak. A platform három fő komponensből nyeri az információkat: a vásárlási rendszerből, amelyeken keresztül a felhasználók megvásárolhatják a játékokat, és ezek a könyvtárukba kerülnek, ezt követően pedig letölthetők; a streaming rendszer, amely biztosítja, hogy bizonyos játékok élő közvetítése elérhetővé váljon, hasonlóan a YouTube vagy Twitch platformokhoz; illetve a közösségi rendszeren keresztül a játékosok hálózatba kapcsolódhatnak, hogy együtt játszhassanak, vagy követhetik egymás tevékenységét. Az ajánlórendszer ezeket az eltérő forrásokat összehangolva generálja az ajánlásokat. Bár a vásárlási, streaming és közösségi platformokon működő ajánlórendszerek már ismertek, ezek egyetlen rend-

¹⁶ Például ezzel foglalkozik Christoph Molnar: *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*. München, 2022.

¹⁷ Thomas Hupperich et al.: An Empirical Study on Price Differentiation Based on System Fingerprints. *arXiv*, 2017.

¹⁸ Gesmann-Nuissl–Meyer i. m. 64–65.

szerbe történő integrálása különleges kihívásokat jelent. Ez a folyamat tanuló, fejlődő és interakcióban lévő neurális hálózatokkal történik, amelyek folyamatosan finomítják az ajánlások pontosságát.¹⁹

Az ajánlórendszerek nem megfelelő tartalmak besorolásával kapcsolatos problémái szintén aggodalomra adnak okot, mivel sértő vagy káros tartalmak téves kategorizálása gyakran előfordult a múltban. Például YouTube 2019-es gyermekvédelmi botránya, amikor az ajánlórendszer olyan tartalmakat javasolt, amelyek potenciálisan veszélyesek voltak gyermekek számára. Az algoritmus a felhasználók érdeklődése és korábbi megtekintései alapján ajánlott videókat, és ennek következtében több esetben felnőttek által feltöltött, ártatlannak tűnő, de szándékosan gyermekeket célzó tartalmak kerültek előtérbe. A probléma különösen akkor vált súlyossá, amikor kiderült, hogy egyes videók alatt a kommentelők olyan időbélyegeket osztottak meg, amelyekben a gyerekek kompromittáló helyzetben voltak, és az algoritmus ezeket a videókat még több felhasználónak ajánlotta. Ez végül a hirdetőik és a hatóságok figyelmét is felkeltette, és a YouTube kénytelen volt módosítani ajánlórendszerét, valamint szigorúbb moderálási szabályokat bevezetni.²⁰

Az online platformok algoritmusai és ajánlórendszerei kulcsszerepet játszanak a visszhangkamrák (*echo chamber*) és szűrőbuborékok (*filter bubble*) kialakulásában is, mivel ezek a rendszerek személyre szabott tartalomajánlásaikkal meghatározzák, hogy a felhasználók milyen információkkal találkozhatnak, ezáltal befolyásolva a közéleti diskurzust és az egyéni véleményformálást. A visszhangkamra és szűrőbuborék fogalmak a digitális hírmédiában az információs izoláció különböző formáira utalnak. A visszhangkamrák esetében a felhasználók tudatosan vagy öntudatlanul olyan közösségekbe zárják magukat, ahol kizárólag saját nézeteiket erősítő információkat fogyasztanak, ezzel fokozva a politikai és társadalmi polarizációt. A szűrőbuborékokat az algoritmus hírszűrés hozza létre, amikor az online platformok személyre szabott tartalmakat jelenítenek meg, így a felhasználók egyoldalú információs környezetbe kerülhetnek, és kevésbé találkoznak eltérő nézőpontokkal.²¹

Az utóbbi évtizedben az ajánlórendszerek társadalmi hatásaira is nagy figyelem irányult, különösen a dezinformáció²² terjedése és a demokratikus folyamatokra jelentett

¹⁹ Uo.

²⁰ The YouTube Team: An update on our efforts to protect minors and families. *YouTube Official Blog*, Jun. 3, 2019. <https://blog.youtube/news-and-events/an-update-on-our-efforts-to-protect/>

²¹ Ezekről részletesen írnak az alábbi szerzők: Papp János Tamás: Ajánlórendszerek és szűrőbuborékok. In: Koltay András – Szikora Tamás – Lapsánszky András (szerk.): *A vadnyugat vége? Tanulmányok az Európai Unió platformszabályozásáról*. Budapest, ORAC, 2024. 231–259.; valamint Gálík Mihály: Visszhangkamrák és szűrőbuborékok. A hírfogyasztásról szóló vita a digitális hálózati média-ökoszisztéma világában. *Médiakutató*, 2020/1.

²² 2025. február 13-án az Európai Bizottság és a Digitális Szolgáltatások Európai Testülete hivatalosan jóváhagyta a dezinformáció elleni magatartási kódex integrálását a DSA keretrendszerébe. A dezinformáció visszaszorítását célzó 2022. évi gyakorlati kódex értelmében: „A dezinformáció olyan bizonyítottan hamis vagy félrevezető információ, amelyet gazdasági haszonszerzés vagy a nyilvánosság szándékos megtévesztése céljából hoznak létre, terjesztenek, és amely közösségi kárt okozhat (például a demokratikus folyamatokra vagy közjavakra – egészség, környezet, biztonság – gyakorolt hatás révén).” Ld. bővebben: Mezei Kitti – Szentgáli-Tóth Boldizsár: Az online platformok használatában rejlő veszélyek:

fenyegetések miatt.²³ Az algoritmusok személyre szabott tartalomkiválasztása révén egyes felhasználók olyan információs buborékba kerülhetnek, amely megerősíti saját nézeteiket, miközben kizárja az eltérő véleményeket. Ez növelheti a társadalmi polarizációt, és manipulációs eszközzé válhat a politikai kommunikációban. Az egyik legismertebb és sokat vitatott eset a Cambridge Analytica-botrány volt 2018-ban,²⁴ amikor politikai kampányok célzottan, személyre szabott üzenetekkel próbálták befolyásolni a választók viselkedését a Facebook algoritmikus hírfolyamán keresztül. Ez az ügy ugyan még a DSA előtt történt, azonban világosan rávilágított arra, milyen mértékű társadalmi és politikai hatással lehetnek az online platformok algoritmusai.

3. A DSA és az algoritmikus átláthatóság

A 2022-ben bevezetett DSA kiemeli az EU jogalkotóinak egyre növekvő aggodalmát az ajánlórendszerek társadalmi hatásaival kapcsolatban. A DSA (70) preambulumbekzdése hangsúlyozza, hogy az ajánlórendszerek jelentős befolyást gyakorolnak a felhasználók online információkeresési és interakciós képességeire azáltal, hogy hozzájárulnak az üzenetek felerősítéséhez, az információk víruszerű terjesztéséhez és az online viselkedés ösztönzéséhez.

A DSA 3. cikk s) pontja szerint ajánlórendszernek minősül minden olyan – teljes mértékben vagy részben – automatizált rendszer, amelyet az online platform arra használ, hogy az online interfészén konkrét információkat javasoljon, vagy ezeket az információkat rangsorolja a szolgáltatás igénybe vevője számára, többek között a szolgáltatás igénybe vevője által indított keresés alapján vagy egyéb módon meghatározva a megjelenített információk relatív sorrendjét vagy elsőbbségét.

A DSA nem csupán az „ajánlórendszerek” jogi fogalmát integrálja az EU szabályozási keretrendszerébe, hanem jelentős mértékben kiterjeszti az algoritmikus átláthatóságra vonatkozó hatályos rendelkezéseket is. Bár az egyes jogszabályok alkalmazási köre eltérhet, az ajánlórendszerek átláthatóságára vonatkozó szabályokat előíró *Platform-to-Business* (P2B) rendelet, a fogyasztói jogokról szóló irányelv (CRD), valamint az irányelv a tisztességtelen kereskedelmi gyakorlatokról (UCPD) elsősorban azon digitális platformokra vonatkoznak, amelyek a platformfelhasználók közötti gazdasági tranzakciók kezdeményezését és lebonyolítását segítik elő.

A DSA hatálya ugyanakkor valamennyi olyan online platformra kiterjed, amely ajánlórendszert alkalmaz, függetlenül attól, hogy az ajánlások konkrét tranzakciók megkönnyítésére irányulnak-e (ezek elsősorban ugyanis adatvezérelt vagy figyelem-piaci szereplők, amelyeknek közös jellemzője a platformalapúság és a zéró áras üzleti

a dezinformáció és a kibertámadások jogi kockázatai. In: Török Bernát – Zódi Zsolt (szerk.): *Az internetes platformok kora*. Budapest, Ludovika, 2022. 319–340.

²³ Ld. erről bővebben: Papp János Tamás: A hamis hírek alkotmányos helyzete és szerepe a demokratikus nyilvánosság befolyásolásában. *In Medias Res*, 2020/1. 141–164.; valamint Falyuna Nóra: Az online dezinformáció és hatásai. *In Medias Res*, 2024/1. 42–74.

²⁴ Margaret Hu: Cambridge Analytica's black box. *Big Data & Society*, Vol. 7., No. 2. (2020) <https://doi.org/10.1177/2053951720938091>

modell).²⁵ Ennek megfelelően az új átláthatósági kötelezettségek olyan platformokra is kiterjednek, mint például a X (korábbi Twitter), a Spotify vagy a Tinder. Az Európai Bizottság eredeti javaslata szerint ezen átláthatósági szabályok kizárólag az online óriásplatformokra vonatkoztak volna, amelyek havi legalább 45 millió aktív felhasználóval rendelkeznek. A trilógus tárgyalások során azonban ezen szabályok alkalmazási körét valamennyi online platformra kiterjesztették, függetlenül a felhasználói kör nagyságától.²⁶

A végleges DSA 27. cikk (1) bekezdése az alábbi kötelezettséget írja elő: minden olyan online platform, amely ajánlórendszert használ, köteles általános szerződési feltételeiben világos és közérthető nyelvezettel feltüntetni az ajánlórendszerek működésének fő paramétereit, valamint azon lehetőségeket, amelyek révén a szolgáltatás igénybevevői módosíthatják vagy befolyásolhatják ezen paramétereket. A DSA 27. cikk (1) bekezdése alapján kötelezi az online platformokat arra, hogy közérthető nyelvezettel ismertessék az általuk használt ajánlórendszerek legfontosabb paramétereit az általános szerződési feltételekben. Az (2) bekezdés tovább részletezi, hogy a legfontosabb kiválasztási szempontokat, valamint azok prioritási sorrendjének okait is meg kell adni. Fontos ugyanakkor hangsúlyozni, hogy az átláthatóság önmagában nem garantálja, hogy a felhasználók érdemben megértsék és kifogásolni tudják az ajánlórendszerek döntéseit. Ehhez további konkrét felhasználói kontrollmechanizmusokra is szükség van. Az ilyen mechanizmusokat a DSA 27. cikk (1) és (3) bekezdése részletezi. Az online platformoknak ismertetniük kell azokat az opciókat, amelyekkel a felhasználók módosíthatják vagy befolyásolhatják az ajánlórendszerek fő paramétereit, ha ilyen lehetőségek rendelkezésre állnak. A (3) bekezdés szerint, ha az ilyen opciók elérhetőek, ezeknek közvetlenül és könnyen hozzáférhetőnek kell lenniük. Fontos megjegyezni, hogy a DSA 27. cikke azonban nem írja elő az ilyen lehetőségek kötelező meglétét, csupán azok hozzáférhetőségét, ha már léteznek. Ugyanakkor a DSA 38. cikke minimumként megköveteli, hogy az online óriásplatformot vagy nagyon népszerű online keresőprogramot üzemeltető szolgáltatók kötelesek a felhasználóknak egy nem személyre szabott ajánlórendszert kínálni.

4. A felhasználói kontroll szintjei

A DSA 27. cikk (1) és (3) bekezdései értelmében a felhasználók az ajánlórendszerek feletti kontrollt három fő szakaszban gyakorolhatják: a bemeneti szint szerint a felhasználók személyes adataik módosításával befolyásolhatják, hogy az ajánlórendszer hogyan nyújtson számukra személyre szabott ajánlásokat; folyamat szinten módosíthatják az algoritmus működését meghatározó paramétereket; a kimeneti szinten megváltoztathatják az ajánlások megjelenítési sorrendjét.

²⁵ Tóth András: A tisztességes adatkereskedelmet biztosító szabályozás szükségességéről. *Állam- és Jogtudomány*, 2021/3. 101–102.

²⁶ Christoph Busch: From Algorithmic Transparency to Algorithmic Choice: European Perspectives on Recommender Systems and Platform Regulation. In: Sergio Genovesi – Katharina Kaesling – Scott Robbins (szerk.): *Recommender Systems: Legal and Ethical Issues*. Springer Verlag, 2023. 46.

4.1. A túlzott információterhelés kockázata

Az ajánlórendszerek „felhasználói ellenőrzési mechanizmusainak”²⁷ megtervezésekor a túlzott információterhelés elkerülése kulcsfontosságú. Egyensúlyt kell teremteni a felhasználóknak biztosított kontroll mértéke és az elfogadható kognitív terhelés között. Ha túl sok technikai információt közölnek az algoritmus működéséről vagy a felhasználói profilról, az túlterhelheti a felhasználókat, ami ellehetetlenítheti a felülvizsgálati lehetőségek gyakorlati alkalmazását. Ezáltal a 27. cikkben megfogalmazott követelmények hatékonysága csökkenhet vagy akár el is veszhet.²⁸

A DSA kifejezetten az ajánlórendszerek felhasználói befolyásolhatóságának kérdésével foglalkozik, az online platformok felhasználóinak megerősítésével. Ennek egyik fő eszköze a már említett átláthatóság, amelyet a 27. cikk rögzít, különös tekintettel a tartalomra és a rangsorolásra. Ezért a DSA 27. cikkének egyik kiemelt célja, hogy a felhasználók megértsék, milyen módon kerülnek bizonyos információk előtérbe számukra, valamint hogyan befolyásolja az online viselkedésük egyes termékek, szolgáltatások vagy tartalmak ajánlását. Ugyanakkor ellentmondás figyelhető meg a 27. cikk által követett átláthatósági célkitűzés és aközött a tény között, hogy a felhasználók gyakran csak felületesen átfutják vagy teljesen kihagyják az online általános szerződési feltételeket, mivel azok gyakran hosszúak és nehezen érthetőek. A DSA előírja, hogy az ajánlórendszereket alkalmazó online platformok szolgáltatói „közérthető és világos nyelvezettel” ismertessék az ajánlórendszereik működését meghatározó fő paramétereket. Ezt az információt a platformok általános szerződési feltételeibe kell beépíteni. Ezen intézkedés célja, hogy a technikai szakszavakkal terhelt algoritmikus folyamatokat érthetővé tegye az átlagos felhasználó számára. A gyakorlati megvalósítás két szempontból is fontos: a jogszabályi megfelelés érdekében ugyanis a komplex algoritmikus folyamatokat közérthető módon kell lefordítaniuk. Az online platformoknak egyértelmű kommunikációs csatornákat kell biztosítaniuk, amelyek révén a felhasználók megérthetik, hogy milyen elvek alapján jelennek meg számukra a tartalmak. Az ajánlórendszerek átláthatóságát a DSA azáltal is előtérbe helyezi, hogy nemcsak a fő paraméterek ismertetését írja elő, hanem azok indoklását is. Ez magában foglalja azokat a kritériumokat, amelyek jelentős hatást gyakorolnak a felhasználók számára megjelenő információkra; az ezek közötti prioritási sorrend meghatározásának okait. Ez egyfajta algoritmikus elszámoltathatóság, amely túlmutat a pusztán átláthatóságon. A szolgáltatóknak nemcsak az alkalmazott változókat kell feltárniuk, hanem azt is, hogy miért és milyen szempontok alapján fontosak ezek a változók. Ez különösen hasz-

²⁷ A tanulmányban az „ellenőrzési mechanizmus” olyan eljárásokat és funkciókat jelöl, amelyek lehetővé teszik a felhasználók számára, hogy az ajánlórendszerek által generált tartalomra vagy ajánlásokra visszajelzést adjanak, azokat vitassák, illetve személyre szabják. E fogalom túlmutat a DSA-ban található „panaszkezelési mechanizmusok” fogalmán, mivel nemcsak jogorvoslatot, hanem a rendszer működésébe való transzparens betekintést és a felhasználói visszacsatolás beépítését is magában foglalja.

²⁸ Christopher Starke – Ljubiša Metikoš – Natali Helberger – Claes de Vreese: Contesting personalized recommender systems: a cross-country analysis of user preferences. *Information, Communication & Society*, Vol. 28., No. 1. (2025) 1–20. <https://doi.org/10.1080/1369118X.2024.2363926>

nos eszközt jelenthet azok számára, akik vizsgálni kívánják, hogy az ajánlórendszerek milyen mértékben erősíthetnek fel rendszerszintű torzításokat vagy félretájékoztatást.

A felhasználói jogok megerősítésének egyik kulcsfontosságú eleme, hogy a 27. és 38. cikkben meghatározott követelményeken kívül az ajánlórendszereket használó online óriásplatformot vagy nagyon népszerű online keresőprogramot üzemeltető szolgáltatók minden ajánlórendszerük esetében legalább egy olyan lehetőséget biztosítsanak, amely nem az (EU) 2016/679 rendelet 4. cikkének 4. pontja szerinti profilalkotáson alapul. A profilozás minden olyan automatizált adatfeldolgozási folyamatot magában foglal, amely egy személy bizonyos jellemzőit elemzi vagy előre jelzi – például viselkedési mintázatait vagy földrajzi helyzetét. Ez a követelmény kiemeli a felhasználói döntéshozatal szerepét, mivel alternatívát kínál a jelenleg domináns, adatvezérelt és gyakran átláthatatlan ajánlórendszerekkel szemben. A DSA tehát egyensúlyt teremt a következő szempontok között: adatvédelem és átláthatóság – a felhasználók nem feltétlenül szeretnék, hogy minden interakciójuk és adatuk befolyásolja az ajánlásokat. Emellett segíti az etikai aggályok kezelését is – az előrejelző algoritmusok gyakran hordoznak rejtett torzításokat, amelyeket így csökkenteni lehet. Ezenkívül ez az előírás technológiai innovációt is ösztönözhet, mivel arra készítheti a vállalatokat, hogy olyan új tartalomajánlási módszereket fejlesszenek ki, amelyek nem támaszkodnak személyes profilozásra.²⁹ A DSA ugyanakkor nem írja elő, hogy az ajánlórendszerek alapértelmezés szerint ne alapuljanak profilalkotáson, jóllehet az Európai Adatvédelmi Felügyelő korábban ezt javasolta a DSA-val összefüggésben.³⁰ Ennek következtében az online platformok alkalmazhatnak olyan ajánlórendszereket, amelyek profilalkotáson alapulnak, és amelyek esetében ez az alapértelmezett beállítás. A DSA alapján az online óriásplatform és a nagyon népszerű online keresőprogram szolgáltatók kötelesek alternatívát biztosítani, azonban a felhasználónak továbbra is aktívan kell kikapcsolnia a profilalkotáson alapuló működést, amennyiben az alapértelmezésként be van állítva az adott platformon. A DSA (94) preambulumbekzdése rögzíti, hogy a profilalkotáson nem alapuló alternatíva kiválasztásának lehetőségét közvetlenül azon online felületen kell biztosítani, ahol a felhasználó az ajánlásokat megtekinti. Ez azt jelenti, hogy az olyan online óriásplatformok esetében, mint például a YouTube vagy a Facebook, az ajánlásokat tartalmazó felületen egyértelmű és jól látható módon – például egy gomb vagy legalább egy legördülő menü formájában – lehetőséget kell biztosítani arra, hogy a felhasználó olyan hírfolyamatot válasszon, amely nem profilalkotáson, hanem például időrendi sorrenden alapul. Műszaki szempontból ez minden további nélkül megvalósítható. Tekintettel arra, hogy számos szolgáltató üzleti modellje jelentős mértékben a profilalkotáson alapuló működésből származó bevételre épül, a DSA által választott opt-out megközelítés kétségkívül egy enyhébb szabályozási megoldásnak tekinthető.³¹

²⁹ Leo Moore – Barry Scannell: For You: AI Recommender Systems and the Digital Services Act. *Lexology*, October 6, 2023. <https://tinyurl.com/3wz26zmp>

³⁰ Opinion 1/2021 on the Proposal for a Digital Services Act. EDPS, 10 February, 2021. <https://tinyurl.com/tbmvuyxx>

³¹ Boris Kandov: Regulatory Approaches or Algorithms on Online Platforms in the Digital Services Act. *ELTE Law Journal*, 2024/2. 136.

Ezek az előírások jelentős mértékben átalakíthatják a felhasználók és az online platformok közötti interakciót, mivel megszüntethetik a felhasználók eddigi passzív szerepét az ajánlórendszerek működésében. A felhasználók ugyanis lehetőséget kapnak arra, hogy módosítsák az ajánlások paramétereit, és ezáltal közvetlenebb módon befolyásolják az általuk látott tartalmakat.

Hogyan hajtják végre a platformok ezeket az új jogi követelményeket? A folyamat jelenleg is zajlik, és a tapasztalatok vegyes képet mutatnak. Jól látható azonban, hogy a platformok eddig elsősorban az átláthatósági követelményekre helyezték a hangsúlyt, miközben a kontrollálhatóság gyakorlati biztosítása továbbra is hiányosnak mondható. Példaként említhető a Meta gyakorlata, ami részletes magyarázatokat vezetett be arról, hogy a Facebook és Instagram ajánlórendszereinek kimenetelei hogyan függenek a különböző tartalomtípusoktól (például Reels, Stories) és az ajánlási mechanizmusoktól (például Explore). Ezek az információk betekintést nyújtanak abba, hogy milyen jelek befolyásolják az ajánlásokat, lehetővé téve a felhasználók számára, hogy megértsék, saját viselkedésük és interakcióik miként befolyásolják az általuk látott tartalmakat.³² A TikTok ezzel szemben egy általánosabb áttekintést nyújt az ajánlórendszereit befolyásoló paraméterekről.³³ Ez az információ azonban elengedhetetlen ahhoz, hogy a felhasználók tudatosabb kontrollt gyakorolhassanak a közösségi médiában megjelenített tartalmak felett.³⁴

4.2. A felhasználói kontroll lehetőségei

A tényleges felhasználói beállítási lehetőségek terén azonban a felhasználók választási lehetőségei jelentősen korlátozottak, mivel a platformok csak szűk körű testreszabást tesznek lehetővé. A Meta platformjain a felhasználók csupán korlátozott számú indokot adhatnak meg arra vonatkozóan, hogy miért nem szeretnének látni egy adott tartalmat. Emellett lehetőségük van bizonyos kulcsszavakat vagy hashtageket megadni, amelyek alapján kiszűrhetik a nem kívánt tartalmakat. A TikTok hasonlóan működik, ugyanis a felhasználók az alkalmazás tartalomeállítási menüjében megadhatnak specifikus szavakat vagy hashtageket, amelyeket szeretnének kizárni a megjelenített tartalmak közül. Ez azonban nem biztosít további testreszabási lehetőségeket, így a Meta megközelítését követi, de nem kínál érdemi plusz opciókat. A Facebook, Instagram és TikTok (DSA 38. cikke alapján) lehetőséget biztosítanak arra is, hogy a felhasználók nem személyre szabott hírfolyamot válasszanak. Ugyanakkor felmerül a kérdés, hogy egy teljesen személyre szabás nélküli élmény mennyire hasznos a felhasználók számára, különösen figyelembe véve a releváns és érdekes tartalmak megjelenítésének szerepét a platformhasználati élményben.

³² Meta Transparency – A rangsoroláshoz használt megközelítésünk. . <https://transparency.meta.com/features/explaining-ranking/>

³³ TikTok Support – How TikTok Recommends Content. <https://tinyurl.com/2ssvrxae>

³⁴ Urbano Reviglio – Matteo Fabbri: The Regulation of Recommender Systems Under the DSA: A Transition from Default to Multiple and Dynamic Controls? *DSA Observatory*, November 22, 2024. <https://tinyurl.com/29f2esz4>

Nyilvánvaló, hogy a DSA 27. cikk alkalmazása eddig elsősorban a felhasználók számára nyújtott átláthatósági követelmények teljesítésére [27. cikk (2)] korlátozódott, míg a felhasználói kontrollmechanizmusok végrehajtása [27. cikk (3)] a legtöbb esetben elmaradt. Úgy tűnik, hogy a platformok egy „minimalista” értelmezést³⁵ követték, amely szerint a 27. cikk nem írja elő a felhasználói beállítási lehetőségek kötelező biztosítását, hanem csupán azok hozzáférhetővé tételét ott, ahol már léteznek.³⁶

4.3. A felhasználói kontroll megvalósításának kihívásai

4.3.1. Felhasználói kontroll az adatok rangsorolási célú felhasználása felett

Az egyik lehetséges megközelítés az, hogy a felhasználók dönthessenek arról, hogy milyen adatok szolgálnak a személyre szabás alapjául. Ez elsősorban azt jelentené, hogy a felhasználók maguk választhatnák meg, hogy profiladataik és implicit preferenciáik (például viselkedési jelek, interakciók) felhasználásra kerüljenek-e az ajánlórendszer által a tartalom személyre szabásához. Az ilyen beállítások lehetőséget adnának a felhasználóknak arra, hogy jobban megértsék, mit „gondol” róluk az ajánlórendszer, és hogy kizárhassák a viselkedési profilozásból fakadó ajánlásokat, amelyek eltérhetnek a tényleges vagy tudatosan kifejezett preferenciáiktól. Másodszor, a felhasználók dönthetnének arról is, hogy milyen preferenciákat vegyen figyelembe az ajánlórendszer a személyre szabás során. Bár számos tervezési stratégia létezik az ajánlások és a felhasználói preferenciák összehangolására, ezt az összhangot hatékonyabban lehetne biztosítani a „személyes címkék” (*personal tags*) operacionalizálásával. Ezek olyan leíró kulcsszavak vagy címkék, amelyek további információt nyújtanak a felhasználó feltételezett érdeklődési köréről. Kína már alkalmazza ezt a megközelítést: a platformok jogszabályi előírás alapján kötelesek biztosítani a felhasználók számára a lehetőséget, hogy kiválasszák vagy eltávolítsák az őket azonosító címkéket. A Douyin, avagy a kínai TikTok, például makrokategóriákba sorolja az érdeklődési területeket, mint például: „Humán tudományok”, „Utazás”, „Gasztronómia”. Ezeken belül további alkategóriák érhetők el. Például a „Gasztronómia” kategória további részekre bontható, mint: „Éttermek felfedezése”, „Különleges fogások élvezete”, „Hagyományos snackek”, „Hozzávalók vásárlása”.³⁷ Amikor egy felhasználó kiválaszt egy kategóriát vagy alkategóriát, lehetősége van meghatározni az érdeklődése prioritását, amely alapján az adott preferencia nagyobb vagy kisebb súllyal esik latba az ajánlórendszer működésében. Bizonyos szempontból ezek a funkciók hasonlítanak a hirdetéskezelési beállításokra, amelyeket például a Google biztosít felhasználóinak.³⁸ Részben ezek a címkék és beállítások megfelelhetnek a DSA által előírt transzparencia követelményeknek, amely szerint a platformoknak közérthetően ismertetniük kell ajánlási kritériumaikat.

³⁵ Starke–Metikoš–Helberger–de Vreese i. m.

³⁶ Reviglio–Fabbri i. m.

³⁷ Shilin Song: Douyin Advertising: a Step by Step Guide. *WalkTheChat*, February 9, 2022. <https://walkthechat.com/douyin-advertising-a-step-by-step-guide/>

³⁸ Google My Ad Center. <https://myadcenter.google.com/home>

A felhasználók dönthetnek arról is, hogy specifikus érdeklődési körök alapján módosítsák az ajánlórendszerek kritériumait. Az Instagramon például a felhasználók korlátozhatják a „érzékeny tartalmak” és a „politikai tartalmak” ajánlásait olyan fiókoktól, amelyeket nem követnek. A TikTok lehetőséget biztosít arra, hogy a felhasználók növeljék a tudomány, technológia, mérnöki tudományok és matematika (STEM) témájú oktatóvideók³⁹ megjelenítését a hírfolyamukban. Bár ezek a funkciók hasznosak lehetnek, „önkéntes módon” kerültek bevezetésre, és nem igazodnak szorosan az EU szabályozás által meghatározott előírásokhoz.⁴⁰ Ezek a funkciók nem jogszabályi előírás alapján, hanem a platformok saját döntése nyomán jelentek meg – például üzleti megfontolásokból vagy a felhasználói elégedettség javítása érdekében. A szolgáltatók maguk határozzák meg, hogy milyen típusú tartalmakra, milyen mértékben biztosítanak szűrési vagy beállítási lehetőséget, ami országonként, célcsoportonként vagy időszakonként eltérhet. Ráadásul a felhasználók sok esetben nem kapnak világos tájékoztatást arról, hogy ezek a beállítások pontosan hogyan működnek.

4.3.2. A felhasználói visszajelzés és annak kihívásai

A felhasználók közvetlen visszajelzése (például kedvelések, elutasítások, „nem érdekel” jelzések egy adott tartalomra vagy témára) kulcsszerepet játszik az ajánlórendszerek működésének befolyásolásában. A DSA azonban ezt a kérdést figyelmen kívül hagyta. A legtöbb esetben ezek a beállítások nehezen hozzáférhetőek, nem eléggé részletesek, vagy csak az alkalmazásban érhetőek el, de a böngészős verzióban nem. Nem világos, hogy a felhasználói visszajelzések hogyan és milyen mértékben befolyásolják az ajánlásokat. Míg a platformok technikailag kiterjeszhetnék a visszajelzési lehetőségeket, továbbra is kérdéses, hogy milyen módon tennék ezt meg. Független auditokra lenne szükség annak ellenőrzésére, hogy ezek az eszközök valóban hatékonyak-e (DSA 37. cikk). A hatékonyság ebben az összefüggésben azt jelenti, hogy a visszajelzések valóban képesek-e érdemben befolyásolni az ajánlórendszer működését, illetve hogy a felhasználók számára érzékelhető-e az eredmény. Ennek megítélése különösen nehéz a felhasználói passzivitás vagy tudatosság hiánya esetén.

5. Az ajánlórendszerek árnyoldalai

2024. október 2-án az Európai Bizottság információkérést nyújtott be a YouTube, Snapchat és TikTok számára a DSA alapján. A megkeresés célja, hogy részletesebb betekintést kapjon az online platformok ajánlórendszereinek felépítésébe és működésébe, különös tekintettel az algoritmusok által alkalmazott paraméterekre. Az Európai Bizottság vizsgálni kívánja, hogy ezek a rendszerek milyen módon erősítenek fel bizonyos rendszerszintű kockázatokat, beleértve a választásokkal és közéleti diskurzusokkal kapcsolatos problémákat, a felhasználók mentális egészségére gyakorolt hatásokat

³⁹ TikTok Support – STEM Feed. <https://support.tiktok.com/en/using-tiktok/exploring-videos/stem-feed>

⁴⁰ Reviglio–Fabbri i. m.

(például függőséget okozó viselkedések és az ún. nyúlüreg hatás⁴¹), valamint a kiskorúak védelmét.

Fontos ugyanakkor elkülöníteni ezt a DSA alapján indult vizsgálatot más, különböző jogalapokon nyugvó eljárásoktól. 2024 márciusában például az olasz hatóság bírságot szabott ki a TikTok számára a UCPD megsértése miatt. Ez az eljárás főként a platform ajánlórendszeréhez kapcsolódó, függőséget kiváltó tervezési elemekre irányult, de túlmutatott az algoritmusok működésén: a határozat tágabb értelemben vizsgálta a platform felelősségét a káros tartalmak terjesztésében, különösen a kiskorúakat és serdülőket érintő kontextusban. A döntés az ilyen tervezési gyakorlatokat agresszívnek minősítette, mivel azok aránytalanul befolyásolják a felhasználókat, és kihasználják egyes felhasználói csoportok – például a fiatalok – sérülékenységét.⁴²

A fentiekkel részben tematikus átfedésben, de eltérő szabályozási logika mentén jelenik meg az Európai Unió mesterséges intelligenciáról szóló rendelete (a továbbiakban: AI Act),⁴³ amely egyes gyakorlatokat kifejezetten tilt. Az AI Act tiltja olyan mesterséges intelligencia rendszerek alkalmazását, amelyek tudatalatti befolyásolási technikákra, célzottan manipulatív vagy megtévesztő eljárásokra épülnek, különösen akkor, ha ezek a felhasználók életkorából, fogyatékosságából vagy társadalmi-gazdasági helyzetéből fakadó sérülékenység kihasználásán keresztül jelentős kárt okoznak, vagy megalapozottan feltételezhető, hogy okoznának.

Az AI Act azt is hangsúlyozza, hogy azokra a gyakorlatokra, amelyek nem tartoznak közvetlenül a tilalmak hatálya alá, más uniós jogszabályok – például a GDPR, a fogyasztóvédelmi szabályok, különösen az UCPD, illetve a DSA – is alkalmazhatók lehetnek. Ezek a szabályozások egymást kiegészítve, komplementer módon járulnak hozzá a felhasználók védelméhez és a digitális környezet biztonságosabbá tételéhez.

A platformoknak ismertetniük kell azokat az intézkedéseket, amelyeket a DSA-ban meghatározott rendszerszintű kockázatok mérséklésére tettek. Ilyen kockázat például a dezinformáció terjedése, a kiskorúakat érintő káros tartalmak megjelenése, a mentális egészségre gyakorolt negatív hatások – például addiktív funkciók vagy algoritmikus torzítások –, valamint az alapvető jogok sérelme, mint a diszkrimináció vagy a véleménynyilvánítás szabadságának korlátozása. Például az algoritmikus torzítás minden ilyen rendszerszintű kockázatban megnyilvánulhat, például egyes vélemények szisztematikus vagy olykor „láthatatlan” eltávolításában (*shadowbanning*, avagy ár-

⁴¹ A nyúlüreg hatás: a felhasználókat egy algoritmus alapján olyan, egymáshoz kapcsolódó, folyamatosan frissülő tartalommal látják el, amelyek ismétlődése és egymásra épülése könnyen függőséghez vezethet. A kifejezés utalás az angolszász nyelvterületen általánosan ismert *Alice's Adventures in Wonderland* meséire, mert a csodaországba Alice a nyúl üregén át jut be.

⁴² AGCM Press Release on Digital Services. Autorità Garante della Concorrenza e del Mercato, 2024. március.

⁴³ Az Európai Parlament és a Tanács (EU) 2024/1689 rendelete (2024. június 13.) a mesterséges intelligenciára vonatkozó harmonizált szabályok megállapításáról, valamint a 300/2008/EK, a 167/2013/EU, a 168/2013/EU, az (EU) 2018/858, az (EU) 2018/1139 és az (EU) 2019/2144 rendelet, továbbá a 2014/90/EU, az (EU) 2016/797 és az (EU) 2020/1828 irányelv módosításáról (a mesterséges intelligenciáról szóló rendelet), HL L, 2024/1689, 2024.7.12.

nyékletiltás).⁴⁴ A DSA (55) preambulumbekzdése továbbá rögzíti, hogy a láthatóság korlátozása történhet még a rangsorolásban, ajánlórendszerekben való hátrébb sorolással, valamint a hozzáférés korlátozásával.

A DSA célja, hogy ezekre a problémákra proaktív kockázatkezelési mechanizmusokat írjon elő az online óriásplatformot vagy nagyon népszerű online keresőprogramot üzemeltető szolgáltató számára. Amennyiben a platformok nem szolgáltatnak megfelelő adatokat ezekről az intézkedéseikről – például ha hiányos, pontatlan vagy félrevezető jelentést nyújtanak be –, az Európai Bizottság formális vizsgálati eljárást indíthat (DSA 66. cikk). Súlyosabb esetekben ez akár bírság kiszabásához is vezethet [DSA 74. cikk (2)]. Ez azt is jelenti, hogy a kockázatkezelési kötelezettség nem csupán formai, hanem érdemi vizsgálat tárgyát képezheti: a hatóságok nemcsak azt nézik, hogy a platform jelentett-e valamit, hanem azt is, hogy a jelentésben szereplő lépések valóságosak és hatékonyak-e.⁴⁵

A DSA 69. cikk (2) bekezdés d) pontja továbbá felhatalmazást ad az Európai Bizottság számára arra, hogy az online óriásplatformot vagy nagyon népszerű online keresőprogramot üzemeltető szolgáltatótól információt kérjen az érintett platform szervezeti felépítéséről, működéséről, informatikai rendszeréről, algoritmusairól, adatkezelési és üzleti gyakorlatáról, valamint ezek magyarázatát is, továbbá, hogy e magyarázatokat rögzítse vagy dokumentálja.

Ez a jogosítvány elsősorban a digitális vállalatok sajátosságaival magyarázható, ugyanis ezek szervezeti struktúrája, működése, különösen pedig technikai infrastruktúrája külső szereplők számára nem, vagy csak nehezen átlátható, nem hozzáférhető, illetve nem értelmezhető. E vonatkozásban a DSA 69. cikk (2) bekezdés b) pontjában rögzített általános vizsgálati jogosítvány sajátos formájának tekinthető, egyfajta digitális vizsgálati jogosultságként értelmezhető. Ez az intézkedés az érintett szolgáltató digitális vagyoni elemeinek átfogó vizsgálatára irányul, ideértve a kiszervezett felhőkapacitásokat és olyan érzékeny információkat is, mint a platform algoritmusai. A vizsgálat helyszíni ellenőrzés formájában történik, de nem korlátozódik a helyben tárolt adatokra. A „magyarázathoz való jog” – és a kapcsolódó dokumentációs kötelezettség – ugyanakkor ennél is tovább megy, figyelembe veszi azt a tényt, hogy a komplex online platformok és keresőmotorok architektúrája még a Bizottság szakértői számára is „fekete dobozként” viselkedhet, ami megnehezítheti, sőt akár ellehetetlenítheti a DSA rendelkezéseinek hatékony végrehajtását. A DSA 69. cikk (2) bekezdés d) pontja tehát, biztosítja a Bizottság számára a kellő szintű technikai átláthatósághoz való jogot a platform vagy keresőmotor működését illetően, amennyiben az a jogséttel összefüggésben releváns. Az ellenőrzés címzettjeinek ezért képesnek és hajlandónak kell lenniük arra, hogy a Bizottság számára érthető és értelmezhető módon bemutassák az

⁴⁴ Corinne Jones: Search Engine Discourse Analysis: How ‘Shadowban’ Affects Policy. *Information Communication & Society*, Vol. 27. (2023) 1025–1042.; valamint Gergely Ferenc Lendvai – Gergely Gosztonyi: Algorithmic Bias as a Core Legal Dilemma in the Age of Artificial Intelligence: Conceptual Basis and the Current State of Regulation. *Laws*, Vol. 14., No. 41. (2025) 7.

⁴⁵ Commission Sends Requests for Information to YouTube, Snapchat, and TikTok on Recommender Systems under the Digital Services Act. European Commission, 2 October, 2024. <https://tinyurl.com/4pzm2x8v>

érintett rendszerek működését – beleértve a szoftverarchitektúrát is, amely jellemzően az ellenőrzés középpontjában áll.

6. Egyéb jogszabályok alkalmazása az ajánlórendszerekre

Az online platformok számára a DSA ajánlórendszerekre vonatkozó szabályai hasonlóságot mutatnak a már meglévő kötelezettségeikkel, amelyeket a P2B rendelet ír elő, és amely 2020-ban lépett hatályba. Míg a DSA azzal foglalkozik, hogy a személyre szabott ajánlások hogyan jönnek létre az online felhasználói viselkedés alapján, addig a P2B rendelet az online platformok által alkalmazott algoritmusokra összpontosít, amelyek meghatározzák az üzleti felhasználók által kínált termékek és szolgáltatások láthatóságát és kiemeltségét.⁴⁶ A P2B rendelet kevésbé az online felhasználói viselkedést veszi alapul, hanem inkább a termékek és szolgáltatások rangsorolására helyezi a hangsúlyt, például az alábbi szempontok alapján: vizuális kiemelések, vásárlói értékelések és vélemények, anyagi ellenszolgáltatás (például fizetett promóciók). Ez a megközelítés üzleti felhasználói szempontból indokolt, mivel az 5. cikk célja, hogy növelje az előrejelezhetőséget az üzleti felhasználók számára. Ennek révén az üzleti szereplők jobban megérthetik az ajánlórendszerek működését, és összehasonlíthatják azokat más online platformokkal. Ez lehetőséget biztosít számukra arra, hogy tudatosan válasszák ki, melyik platformon kívánják termékeiket vagy szolgáltatásaikat értékesíteni. Ha az üzleti felhasználók tisztában vannak azzal, hogyan kerülnek rangsorolásra a termékek és szolgáltatások, akkor javíthatják saját megjelenítésüket, ezáltal növelhetik üzleti sikerüket. Ahogyan a DSA, úgy a P2B rendelet is előírja, hogy az online platformoknak általános szerződési feltételeikben ismertetniük kell, hogy az üzleti felhasználók milyen módon tudják aktívan befolyásolni a rangsorolást. Azonban van egy lényeges különbség: a P2B rendelet csak azt szabályozza, hogy az üzleti felhasználók pénzügyi ellenszolgáltatás révén hogyan tudják befolyásolni a rangsorolást. Ezzel szemben a DSA egy ennél szélesebb jogot biztosít: lehetőséget ad az online preferenciák módosítására is, nemcsak a rangsorolás befolyásolására fizetett promóciókkal. Ez azt jelenti, hogy míg a P2B rendelet főként az üzleti szereplők kiszámíthatóbb működését és a piaci versenyt segíti elő, a DSA elsődleges célja a végfelhasználók jogainak erősítése és az ajánlórendszerek feletti kontroll biztosítása.⁴⁷ Az P2B rendelet szerint a rangsorolás az online közvetítő szolgáltatásokon kínált áruk és szolgáltatások relatív kiemelését és strukturálását jelenti. A DMA⁴⁸ is tartalmaz a rangsorolásra vonatkozó szabályokat, de ezek elsősorban versenyjogi megfontolásokat tükröznek: a kapuőröknek átlátható, tisztességes és megkülönböztetésmentes feltételeket kell biztosítaniuk, és nem részesíthetik előnyben saját szolgáltatásaikat. Ezzel szemben a DSA nem gazdasági, hanem

⁴⁶ Ld. erről Zódi Zsolt: Platformjog és felhasználóvédelem. *Gazdaság és Jog*, 2022/3–4. 15–18.

⁴⁷ Anna-Lena Kempf – Michelle Seel: How the Digital Services Act Changes Things for Platforms. *Pinsent Masons*, 20 Feb., 2024. <https://tinyurl.com/4w43umyu>

⁴⁸ Az Európai Parlament és a Tanács (EU) 2022/1925 rendelete (2022. szeptember 14.) a digitális ágazat vonatkozásában a versengő és tisztességes piacokról, valamint az (EU) 2019/1937 és az (EU) 2020/1828 irányelv módosításáról (digitális piacokról szóló jogszabály). HL L 265., 2022.10.12., pp. 1–66

társadalmi és közérdekű szempontok alapján szabályozza az ajánlórendszerek működését, így azokra szélesebb körű és részletesebb követelményeket határoz meg, különösen az átláthatóság, felhasználói autonómia és kockázatkezelés területén. A DMA (52) preambulumbekzdése hangsúlyozza, hogy a rangsorolás fogalma „minden relatív kiemelési formát magában foglal, beleértve a megjelenítést, értékelést, összekapcsolást vagy hangalapú eredményeket.”⁴⁹ Ebben az összefüggésben a DMA egy érdekes pontositást tesz, miszerint a rangsorolás magában foglalhatja azokat az eseteket is, amikor egy alapvető platformszolgáltatás kizárólag egyetlen eredményt jelenít meg vagy közvetít a végfelhasználó felé. Első ránézésre ez ellentmondásosnak tűnhet, azonban valószínűleg az az oka, hogy a DMA által szabályozott „alapvető platformszolgáltatások” közé tartoznak a virtuális asszisztensek is. Amikor a fogyasztók virtuális asszisztenseket használnak online vásárlásra, az Amazon Alexa és más hang vezérelt eszközök nem olvasnak fel hosszú, rangsorolt terméklistákat, hanem csupán egyetlen ajánlatot jelenítenek meg (mint az „*Amazon's Choice*” ajánlat). Ez bizonyos értelemben a rangsorolás fogalmának egy szélsőséges alkalmazását jelenti. A DSA más terminológiát alkalmaz, és a „*recommender system*” (ajánlórendszer) technikai kifejezést használja. A különböző megfogalmazás ellenére a szabályozás tartalma lényegében megegyezik a többi jogszabályszöveggel, amelyek a „rangsorolás” fogalmát alkalmazzák, mint például a P2B és a DMA is. A DSA célja sem különbözik más jogi eszközöktől, ugyanis azon rendszerek szabályozását célozza, amelyek meghatározzák a „relatív kiemelés sorrendjét” és így döntési architektúrát alakítanak ki a platformfelhasználók számára. Mindezek alapján megállapítható, hogy a különböző terminológia ellenére a releváns jogszabályok egy fontos közös jellemzővel rendelkeznek: nem tesznek különbséget az ajánlórendszerek típusai között (például tartalom alapú, tudásalapú, együttműködésen alapuló vagy hibrid rendszerek). Ehelyett technológiásemleges megközelítést alkalmaznak, és kizárólag az ajánlórendszerek kimenetére összpontosítanak, azaz arra, hogy milyen „relatív sorrendet” vagy „kiemelést” hoznak létre.⁵⁰

7. Kockázatértékelés szerepe az átláthatóságban

A szolgáltatóknak átfogó kockázatértékeléseket kell végezniük, különös tekintettel a platformjaikon jelentkező rendszerszintű kockázatokra. Ennek részeként elemeznük kell, hogy az ajánlórendszerek és egyéb algoritmusok hogyan befolyásolják ezeket a kockázatokat. Ez különösen fontos a dezinformáció és káros tartalmak terjedésének vizsgálatakor, amelyeket az algoritmusok akár felerősíthetnek. A DSA előírja, hogy a

⁴⁹ DMA 2. cikk 22. „rangsorolás: az online közvetítő szolgáltatásokon, online közösségi hálózati szolgáltatásokon, videomegosztóplatform-szolgáltatásokon vagy virtuális asszisztenseken keresztül kínált áruknak és szolgáltatásoknak biztosított relatív elsőbbség, illetve az online keresőprogramok által a keresési eredményekhez rendelt relevancia, ahogyan azokat az online közvetítő szolgáltatást nyújtó vállalkozások, az online közösségi hálózati szolgáltatások, a videomegosztóplatform-szolgáltatások, virtuális asszisztensek vagy az online keresőprogramok megjelenítik, strukturálják vagy kommunikálják, függetlenül az ilyen megjelenítéshez, strukturáláshoz vagy kommunikációhoz használt technológiai eszközöktől, valamint attól, hogy csak egyetlen eredményt jelenítenek-e meg vagy kommunikálnak-e.”

⁵⁰ Busch i. m. 36–37.

platformoknak fel kell tárniuk a szolgáltatás szándékos manipulációjából eredő kockázatokat, például botok vagy csalárd tevékenységek révén történő dezinformációt. Ez biztosítja, hogy ne csak az algoritmusok hibái, hanem az azokkal kapcsolatos rosszindulatú visszaélések is megfelelő figyelmet kapjanak az online platformok részéről. Az online óriásplatformok és a nagyon népszerű online keresőmotor szolgáltatók évente kötelesek jelentést készíteni a rendszerszintű kockázatok csökkentése érdekében tett lépéseikről, beleértve az ajánlórendszereik módosítását és azok hatását. Ezzel összefüggésben érdemes megemlíteni, hogy a legnagyobb platformok – köztük a Meta és a TikTok – már közzétették első jelentéseiket a DSA 34–35. cikke alapján. Ezekből kiderül, hogy milyen konkrét lépéseket tettek az ajánlórendszerek átalakítására, például a dezinformációs tartalmak terjedésének visszaszorítására vagy a nem profilalapú ajánlások elérhetővé tételére. A jelentések ugyanakkor jelentős eltéréseket mutatnak abban, hogy a platformok milyen részletességgel számolnak be ezek hatásairól – ami további szabályozási kérdéseket is felvet. A Meta 2024. évi DSA kockázatkezelési jelentése szerint a Facebook és az Instagram algoritmusait módosították annak érdekében, hogy azok „ne járuljanak hozzá a gyorsan, víruszerűen terjedő dezinformáció automatikus felerősítéséhez”.⁵¹ Ez arra utal, hogy a platform tudatosan próbálja csökkenteni az olyan tartalmak elérését, amelyek gyorsan terjednek, de alacsony megbízhatóságú forrásból származnak. Ugyanakkor a jelentés nem részletezi, hogy pontosan milyen mechanizmusokon keresztül történik ez a szűrés, illetve hogy a módosítások mennyire befolyásolják a felhasználói élményt vagy a tartalom sokféleségét. Ez felveti az átláthatóság és az elszámoltathatóság kérdését: a felhasználók és szabályozók számára továbbra sem világos, milyen konkrét beavatkozások történnek az algoritmus szintjén, és azok milyen eredményre vezetnek.

A TikTok jelentésében ezzel szemben egy proaktív tartalomirányítási példát emelnek ki: a STEM (tudomány, technológia, mérnöki tudományok, matematika) témájú tartalmakra létrehozott dedikált hírfolyamot. A „*STEM feed*” célja, hogy növelje a megbízható, oktató jellegű videók arányát a felhasználói ajánlások között, és ezzel alternatívát kínáljon a káros vagy dezinformációs tartalmakkal szemben.⁵² Ez a megközelítés pozitív ösztönzést jelent a tartalomajánlás szintjén, és jól illusztrálja, hogy az algoritmikus tervezés nemcsak korlátozásra, hanem tartalomkiemelésre is alkalmas lehet. Ugyanakkor itt is felmerül a kérdés, hogy milyen algoritmikus logika szerint valógot a rendszer, és hogy valóban eléri-e a kívánt célközönséget.

8. Összefoglalás: Út a valódi átláthatóság felé?

A DSA történelmi változást hozhat az online platformok felhasználói élményében, de ennek eléréséhez szükség lenne egyértelmű iránymutatásokra, magatartási kódexre vagy akár felhatalmazáson alapuló jogi aktusokra, amelyek részletezik, hogyan kell biztosítani az ajánlórendszerek átláthatóságát és felhasználói kontrollját.

⁵¹ Meta's Progress Implementing the Digital Services Act. Meta, November 28, 2024. <https://tinyurl.com/3p8hxx34>

⁵² TikTok: Digital Services Act Reporting. <https://www.tiktok.com/transparency/en/dsa-transparency>

A DSA által előírt átláthatósági kötelezettségek – különösen a 27. és 38. cikk – fontos újtásként jelennek meg, mivel előírják, hogy az online platformok közérthető formában tájékoztassák a felhasználókat az általuk alkalmazott ajánlórendszerek fő paramétereiről, illetve kínáljanak lehetőséget nem profilalapú ajánlások választására. Ugyanakkor ezek a szabályozások – bár előrelépést jelentenek – nem biztosítanak tényleges felhasználói kontrollt, különösen annak fényében, hogy a DSA nem írja elő kötelező jelleggel az alternatív beállítások biztosítását, csupán azok elérhetőségét ott, ahol már léteznek.

Az ajánlórendszerek működéséről szóló magyarázatok ugyan hasznosak lehetnek, önmagukban nem elegendők, ha ezek nem társulnak érdemi felhasználói kontrollal. Az algoritmusok működéséről adott technikai magyarázatok gyakran túlságosan összetettek a laikus felhasználók számára, így azok nem segítik elő a tudatos döntéshozatalt. Ezért szükség lenne egyszerűbb, interaktívabb és személyre szabható mechanizmusokra, amelyek révén a felhasználók valós befolyást gyakorolhatnának arra, hogy milyen szempontok alapján kapnak ajánlásokat. A szabályozás hatékonysága csak akkor érvényesülhet, ha a felhasználók számára nemcsak információt biztosítanak az algoritmusok működéséről, hanem gyakorlati, közérthető és hozzáférhető eszközöket is nyújtanak az ajánlási folyamat alakításához. A jelenlegi online platformok által bevezetett átláthatósági és felhasználói kontrollmechanizmusok egyelőre csak „látszatátláthatóságot” (*transparency washing*)⁵³ eredményeznek, valódi befolyásolási lehetőség nélkül.⁵⁴ Éppen ezért a jogalkotóknak a jövőbeli jogi szabályozás továbbfejlesztésekor szem előtt kell tartaniuk a technológiai, társadalmi és pszichológiai tényezőket egyaránt, mivel az algoritmikus rendszerek hatása komplex módon jelenik meg az online térben.

⁵³ Monika Zalnieriute: “Transparency-Washing” in the Digital Age: A Corporate Agenda of Procedural Fetishism. *Critical Analysis of Law*, Vol. 8., No. 1. (2021) 39–53. UNSW Law Research Paper No. 21-33. Available at SSRN: <https://ssrn.com/abstract=3805492>

⁵⁴ Reviglio–Fabbri i. m.

ÚJ KORSZAKOT NYIT-E A MESTERSÉGES INTELLIGENCIA A DEZINFORMÁCIÓBAN?

*Az MI szerepének értelmezése a dezinformációs lánc
mint új elméleti keret segítségével*

Pintér Róbert*

1. Bevezetés

1.1. Hihetünk-e a fülünknek és a szemünknek?

A World Economic Forum 2024-es globális kockázati jelentése szerint a mesterséges intelligencia (MI) által generált dezinformációt és félretájékoztatást (*misinformation*) tartották az emberek a második legfontosabb olyan tényezőnek, ami világméretű válsághelyzetet idézhet elő a következő két évben. Első helyre a szélsőséges időjárás, harmadikra pedig a társadalmi és/vagy politikai polarizációt tették a kutatásban részt vevők.¹ Arról, hogy mit jelenthet egy ilyen dezinformáció okozta válsághelyzet, Martin Ford már 2022-ben érzékletesen írt az MI-ről szóló könyvében: képzeljük el, hogy két nappal az amerikai elnökválasztás előtt kiszivárog egy valódinak tűnő generált hangfelvétel az egyik jelöltről, amiben nevetve magyarázza, hogy mindig is fajgyűlölő volt:

„[m]iközben a felvételtől felrobban az internet, és a televíziócsatornák is újra meg újra leadják, a közösségi média világában eluralkodik a káosz és a felháborodás. A politikus ádáz harcok árán jutott el a jelölésig, és feldühödött ellenfelei most a visszalépését követelik.”²

* Egyetemi docens, Budapesti Corvinus Egyetem, Adatelemzés és Informatika Intézet.
ORCID: <https://orcid.org/0000-0001-7807-7997>

¹ The Global Risks Report. World Economic Forum (WEF), 2024.
https://www3.weforum.org/docs/WEF_The_Global_Risks_Report_2024.pdf

² Martin Ford: *Jövönk a robotok korában – Hogyan alakítja a mesterséges intelligencia az életünket?* Budapest, HVG, 2022. 207.



Sokan attól tartottak, hogy a 2024-es, világszerte nagyjából 2 milliárd embert és 60 országot érintő választásokon a mesterséges intelligencia segítségével generált dezinformáció a fentihez hasonló módon érdemben befolyásolhatja majd a szavazói döntéseket és ezzel aláássa a demokratikus működést.³ Az amerikai kongresszus 2023-ban törvényjavaslatot is előterjesztett, amely előírja az MI által generált tartalmak egyértelmű jelölését a politikai hirdetésekben.⁴ Ennek hatására a Meta (a Facebook anyavállalata) és más platformok is kötelezővé tették az MI-vel készített vagy módosított tartalmak megjelölését. Az intézkedések a dezinformáció veszélyeire reagáltak, különösen a *deepfake*⁵ és a személyre szabott politikai hirdetések kapcsán.⁶ A Pew Research Centre 2024 augusztusi felmérése szerint az amerikaiak 39%-a aggódott az MI választásokra gyakorolt potenciálisan káros hatásai miatt, ami mind a két politikai oldal szavazóira egyaránt jellemző volt.⁷

Egy-két, nagyobb hírverést keltő esetet leszámítva⁸ azonban az MI által generált tartalmak felhasználása szerencsére jóval alatta maradt a várakozásoknak és nem igazolta az előzetes félelmeket. A romániai elnökválasztás megsemmisített első fordulójának kivételével mindenhol rendben lezajlottak a 2024-es választások. Egyébként a romániai esetben sem az MI-vel generált tartalmak okozták a problémát, hanem a 2008-as Obama kampány óta⁹ politikai kommunikációs szempontból immár hagyományosnak tekinthető közösségimédia-platformok használata, elsősorban a TikTok-on zajló összehangolt dezinformációs kampány.¹⁰

Ennek ellenére a generatív MI-vel kapcsolatos félelmek továbbra is meghatározóak és az aggodalmak eljutnak a nagyközönséghez is, mivel olyan, széles körben ismert szakemberek is fenntartásuknak adnak hangot, akikre világszerte odafigyelnek. Sir Tim Berners-Lee, a WWW atyja egy interjújában kijelentette, hogy „sokkoló, mérgező

³ Ld. például Marina Adams: How AI-generated disinformation might impact this year's elections and how journalists should report on it. *Reuters Institute*, 2024. március 15. <https://tinyurl.com/yc3rrxtz>; vagy Vittoria Elliott: 2024 Is the Year of the Generative AI Election. *Wired*, 2024. május 30. <https://www.wired.com/story/2024-is-the-year-of-generative-ai-elections/>

⁴ 118th US Congress. S. 1596, 2023. május 13. <https://tinyurl.com/pxnvs4zm>

⁵ MI-vel szerkesztett vagy generált képek, videók és hangfelvételek, amelyek valós vagy nem létező embereket ábrázolnak.

⁶ Social Media, Disinformation, and AI: Transforming the Landscape of the 2024 U.S. Presidential Political Campaigns. *The SAIS Review of International Affairs*, 2025. január 14. <https://tinyurl.com/muuv722u>

⁷ Shanay Gracia: Americans in both parties are concerned over the impact of AI on the 2024 presidential campaign. *Pew Research Center*, 2024. szeptember 19. <https://tinyurl.com/wjemsv3r>

⁸ Ilyen jól dokumentált esetek voltak, amikor Biden elnök állítólag nem-szavazásra buzdított telefonon az előválasztáson, vagy egy már halott, de MI-vel feltámasztott indiai politikus támogatásáról biztosította az egyik jelöltet, illetve hírességek, például Eminem vagy Taylor Swift is látszólag politikai állást foglaltak és nem csak az amerikai választásokon. További példák elérhetőek a *Wired* magazin gyűjteményében arról, hogy hogyan használták a generatív MI-t a 2024-es választásokon: Vittoria Elliott: The Wired AI Elections Project. *Wired*, May 30, 2024. <https://www.wired.com/story/generative-ai-global-elections/>

⁹ Bruce Bimber: Digital Media in the Obama Campaigns of 2008 and 2012: Adaptation to the Personalized Political Communication Environment. *Journal of Information Technology & Politics*, Vol. 11., No. 2. (2014) 130–150. <https://doi.org/10.1080/19331681.2014.895691>

¹⁰ Théophane Hartmann: Commission opens TikTok investigation over Romanian presidential elections disinformation. *Euractiv*, 2024. december 17. <https://tinyurl.com/y8shvk3y>

algoritmusok” működnek a közösségi médiában és a megoldás a programozók kezében van. Szerinte a világháló „a dezinformáció, a polarizáció és a vállalati visszaélés játszótérévé változik”.¹¹ Az izraeli sztártörténész, Yuval Noah Harari szerint az MI a demokráciához szükséges társadalmi szintű párbeszédet lehetetleníti el. Szerinte a demokratikus berendezkedés mindaddig kisebb közösségekre (törzsekre, városállamokra) szorítkozott, amíg nem volt meg a technológiai lehetősége a társadalmi szintű párbeszédnek (vagyis a sajtó megjelenése előtt). A modern demokrácia alapja a párbeszéd, ha ez az MI miatt ellehetetlenül, az a demokrácia végét jelentheti.¹²

A dezinformáció és az MI kapcsolata nem csak külföldön, hanem Magyarországon is széles tudományos érdeklődésre tart számot. Nemrégiben például konferenciát is szenteltek a témának.¹³ A Mérték Médiaelemző Műhely 2023-ban dezinformációról szóló tanulmánygyűjteményt adott ki.¹⁴ Aczél Petra és Veszelszki Ágnes szerkesztésében a *deepfake*-ről jelent meg könyv számos szerző közreműködésével.¹⁵ Krekó Péter összeesküvés elméletekről és dezinformációról szóló könyve pedig több kiadást is megért, ami a téma iránti szélesebb, laikus közönség érdeklődését is jól mutatja.¹⁶ Krekó és munkatársai *Dezinformáció és Mesterséges Intelligencia Kutatócsoport* néven önálló kutatócsoportot is felállítottak a téma pszichológiai vetületeinek vizsgálatára az ELTE-n.¹⁷

Az MI-vel kapcsolatos aggodalmak a generatív MI, elsősorban a ChatGPT 2022 végi berobbanásával és használatának széles körben való elterjedésével erősödtek fel. A *deepfake* tartalmak kapcsán egyre gyakrabban hangoztatott állítás, hogy ezentúl ‘nem hihetünk a szemünknek és a fülünknek’,¹⁸ ami egyúttal azt is sugallja, hogy új korszakot hoz az MI a dezinformációban. Jelen tanulmányomban azt vizsgálom meg, hogy az MI milyen szerepet játszik a dezinformációban és vajon tényleg új korszakot jelenthet-e.¹⁹

¹¹ Sokkoló, mérgező algoritmusokra figyelmeztet a világháló atyja. *HVG*, 2024. december 8.

https://hvg.hu/tudomany/20241208_tim-berners-lee-interju-kozossegi-media-veszelyei

¹² Yuval Noah Harari: *Nexus – Az információs hálózatok rövid története a kőkorszaktól az MI-ig*. Budapest, Animus, 2024.

¹³ Nóra Falyuna – Katalin Feher – Márton Demeter – Gábor Szüdi – Joseph E. Uscinski: Against Disinformation: Bridging Science and Public Discourse. *Információs Társadalom* Vol. 24., No 2. (2024) 68–84. <https://dx.doi.org/10.22503/inftars.XXIV.2024.2.4>

¹⁴ Çakır-Somlyai Özge (szerk.): *A dezinformáció jelensége – tanulmánygyűjtemény*. (Mérték Füzetek 33.) Budapest, Mérték Médiaelemző Műhely, 2023. <https://tinyurl.com/dfpurnw3>

¹⁵ Aczél Petra – Veszelszki Ágnes (szerk.): *Deepfake: a valótlán valóság*. Budapest, Gondolat, 2023.

¹⁶ Krekó Péter: *Tömegparanoia 2.0 – Összeesküvés-elméletek, álhírek és dezinformáció*. Budapest, Athaneum, 2023.

¹⁷ A kutatócsoportról bővebben ld. <https://tinyurl.com/3xv38xjz>

¹⁸ Ld. pl. Nagy Krisztina: Hihetünk-e még a szemünknek? Mit és hogyan tanítsunk a dezinformáció és a mesterséges intelligencia kapcsolatáról az Alfa és a Z generációnak? Konferencia előadás. *Mesterséges intelligencia – filozófia, jog és dezinformáció konferencia*, Budapest, 2024. november 20. <https://tinyurl.com/4mzcm9nw>

¹⁹ A tanulmányban nem foglalkozom érdemben a dezinformáció és az MI szabályozási vonatkozásaival, bár érintőlegesen több ponton is előkerül a kérdés. Akit ez a téma mélyebben érdekel, annak a következő cikket javaslom kiindulópontnak: Bilge Azgin – Sevki Kıralp: Surveillance, Disinformation, and Legislative Measures in the 21st Century: AI, Social Media, and the Future of Democracies. *Social Sciences*, Vol. 13., No. 10. (2024) 510. <https://doi.org/10.3390/socsci13100510>

Az egészen biztos, hogy sok mindenre képes a generatív MI, ami korábban nem volt lehetséges, legfeljebb csak nagy költségek és erőforrások mozgósítása révén. A generatív MI alkalmas szövegek, képek, mozgóképek és hangfelvételek létrehozására és módosítására. Ez a képesség az utóbbi években bárki számára hozzáférhetővé vált az interneten elérhető generatív MI eszközök (elsősorban a ChatGPT, illetve az olyan képgeneráló alkalmazások, mint a DALL-E vagy a Midjourney és például az ElevenLabs fémjelzett hanggeneráló szolgáltatások) elterjedésével. Csakhogy a technológia képességének megváltozása miatt önmagában még nem feltétlenül köszönt be új éra a dezinformációban. Hiszen elsősorban nem azért *nem csapjuk be egymást*, mert ne lennénk képesek rá (vagyis korábban ne lettek volna meg az eszközeink hozzá), hanem mert a viselkedési normák kölcsönös betartása és a bizalom az alapja az emberi társadalmak működésének. A kérdés – új korszakot hoz-e az MI a dezinformációban – vizsgálata tehát elsősorban nem az MI képességeiről szól, hanem azok emberi felhasználásáról, illetve ennek következtében a minket körülvevő világ működésének esetleges megváltozásáról.

1.2. A dezinformáció fogalma és a dezinformációs lánc

Írásomban a dezinformáció médiatudományi definíciójánál tágabban értelmezem a fogalmat. „Az Európai Bizottság Álhír- és Online Dezinformációs Csoportjának 2018-as definíciója szerint a dezinformáció olyan »hamis, pontatlan vagy félrevezető információ, amelyet a nyilvánosságnak való károkozás vagy a profitszerzés szándékával terveznek meg, mutatnak be és terjesztenek«.”²⁰ Meglátásom szerint ugyanis, mint látni fogjuk, a motivációk ennél jóval tágabbak lehetnek, miközben nem minden dezinformációt szánunk a széles nyilvánosságnak – gondoljunk csak az unokázós csalásokra. Az én értelmezésemben a dezinformáció olyan információ, amit függetlenül valóságtartalmának mértékétől, a terjesztés formájától, közlésének módjától, címzettjétől, befogadásától és ténylegesen elért hatásától, azzal a céllal hoz létre valaki, hogy azt szándékosan félrevezetésre használja. Így nem csak a hírek esnek bele ebbe a fogalomba, ez alapján kerülöm is az ‘álhír’ mint kifejezés általános használatát, miután az túlságosan leszűkítő. De a Barts Lívia által meghatározott dezinformáció meghatározást – „[d]ezinformációról akkor beszélünk, amikor hamis vagy megtévesztő információt osztanak meg tudatosan, hogy kárt okozzanak”²¹ – sem tartom teljesen elfogadhatónak. Ugyanis, mint látni fogjuk, a dezinformáció mögött sokkal színesebb motivációkat találni, mint pusztán a *károkozás*, így a dezinformáció jelentős részét kizárnánk a vizsgálódásból, ha elfogadnánk ezt a meghatározást. A legközelebb Z. Karvalics László megközelítése áll hozzám: „~szándékosan felépített megtévesztő, félrevezető tartalom; a valóságnak a kiadó számára előnyös értelmezése vagy meghamisítása, amelyet a befogadó számára megalapozott és ellenőrzött információként tárnak, és a hatás elérése érdekében hihetővé tesznek.”²²

²⁰ Bajomi-Lázár Péter: Álhírek és igaz hírek – Definíciós kérdések. In: Çakır-Somlyai (szerk.) i. m. 8.

²¹ Barts Lívia: Vizuális információs zavarok – típusok és módszerek. In: Çakır-Somlyai (szerk.) i. m. 35.

²² Z. Karvalics László: *Informatorium – Word Guide to Contemporary Information Culture*. Budapest, Tinta Publishing House, 2017. 142. (saját fordítás P.R.)

A dezinformáció fogalompárja a miszinformáció (félretájékoztatás vagy téves információ). Az én értelmezésemben ez olyan dezinformáció, ahol a terjesztésben részt vevő nem tud arról, hogy dezinformációt terjeszt. Vagyis ugyanaz az információ a dezinformáción túl miszinformáció is lehet a detektálástól, a befogadástól és a terjesztés módjától függően.

A cikkben a kutatási kérdés vizsgálatának részeként javaslatot teszek a dezinformáció egyfajta új értelmezési keretére, amit *dezinformációs láncnak* neveztem el. Azt gondolom ugyanis, hogy a dezinformáció szisztematikus vizsgálata segíthet annak eldöntésében, vajon újat hozhat-e a dezinformációban a mesterséges intelligencia. A dezinformációs lánc gondolata gyakorlatilag annak a nyolc hónapos kutatómunkának az eredménye, amit a cikk elkészítése igénybe vett. A fogalmat a szakirodalmak olvasása és a dezinformációs jelenségek értelmezése közben alakítottam ki, annak felismerése nyomán, hogy ezek a jelenségek felfűzhetők egyetlen láncra, ami támogatja azok kapcsolatának jobb megértését.

A következő fejezetekben tehát a dezinformációs lánc hat kulcs lépésén megyek végig és vizsgálom meg azokat egyenként, hogy jelenleg mit tudunk róluk és milyen változások érhetők tetten bennük az MI-vel való kapcsolatuk alapján. A dezinformációs lánc főbb lépései (a lépések kezdőbetűit összeolvasva MoLTe(De)BeHa):

1. Motiváció: a dezinformációt folytató motivációjának azonosítása
2. Létrehozás: a dezinformációs üzenet létrehozásának aktusa
3. Terjesztés: a dezinformáció terjesztésének folyamata
4. (Detektálás: a dezinformáció tényének detektálása – amely nem minden esetben történik meg, ezért itt zárójelben szerepel)
5. Befogadás: a célszemély(ek) részéről a dezinformáció percepciója, feldolgozása és befogadása
6. Hatás: végül a dezinformációval elért hatás, illetve a szélesebb (nem szándékolt) társadalmi következmények.

Mint később látni fogjuk ez a fajta megközelítés jelentősen kibővíti Wardle és Derekshan információs zavarokról szóló leírását (aminek része a dezinformáció is), ahol a szerzők három fázist különítettek el az információs zavarok vonatkozásában, ezek a létrehozás (az üzenet létrehozása) – (újra)termelés (üzenet, mint média termék) – terjesztés (üzenet nyilvános megosztása).²³

2. A dezinformációs lánc hat kulcs lépése

A jelenség súlyának ismertetése, a központi kutatási kérdés bemutatása és a fogalmi tisztázás után vegyük végig a dezinformációs lánc egyes lépéseit és vizsgáljuk meg abban az MI lehetséges szerepét, különös tekintettel a generatív MI-re.

²³ Claire Wardle – Hossein Derakhshan: *Information Disorder: Toward an Interdisciplinary Framework for Research and Policymaking*. Strasbourg, Council of Europe Report, 2017. <https://tinyurl.com/4z7se79j>

2.1. Első lépés: motiváció

Bár, mint fentebb láttuk, a szakirodalom hajlamos elsősorban politikai motivációkra visszavezetni a dezinformációt, valójában még a politikai hatással bíró dezinformáció mögött is lehet egészen más indíték. Például a 2016-os amerikai elnökválasztás alatt macedón fiatalok kiterjedt dezinformációs hálózatot üzemeltettek, ami mögött pusztán a meggazdagodás vágya állt. A jól dokumentált esetben az autodidakta, angolul is rosszul tudó ‘vállalkozók’ azt használták ki, hogy az online hirdetési platformok az internetes forgalom mérete alapján fizetnek a hirdetést futtató oldalaknak.²⁴ Mivel az amerikai választáshoz kötődő álhírekkel nagy forgalmat lehetett generálni, ezért dezinformációs oldalakat hoztak létre, amelyek óriási látogatottságot produkáltak. A történetben a veszei fiatalok szemfülességén és etikátlanságán túl legalább ennyire fontos az interneten bevett reklám alapú üzleti modell szerepe és azon cégek felelőssége, amelyek nem tiltják a dezinformációra alkalmas tartalom mellett hirdetések futtatását. Nagy valószínűséggel persze ez az eset (gazdasági okokból felépülő politikai dezinformációs gépezet) inkább kivételnek, semmint általánosnak tekinthető. De a célzott, akár személyre szabott dezinformációs aktivitások mögött is sokszor gazdasági haszonszerzés áll (ilyenek például a *deepfake* hirdetések, a pénzhamisítás, a festmények hamisítása vagy az unokázós csalás).

Természetesen a legtöbb, nagy hatással bíró dezinformáció mögött politikai motivációt, zavarkeltést, vagy egyenesen a demokratikus intézmények működésének aláásását találni. Ugyancsak ide sorolható a zaj és információs túltelítettség, tényrelativista helyzet létrehozása. Ezeknél a cél, hogy „olyan »fehér zajt« – információs túltelítettséget – teremtsenek, amelyben nehéz megkülönböztetni az igaz és a hamis állításokat”.²⁵ Miután a cikkben később a politikai indíttatású dezinformáció áll majd az elemzés középpontjában, így erről itt nem írnék bővebben.

Egy további fontos motiváció lehet a dezinformációra a szórakoztatás és a humor. Ezek egyik legismertebb formája a *mockumentary* (fiktív eseményeket ábrázoló, de dokumentumfilmként bemutatott műsor, például reality sorozatok) és az áldokumentumfilmek. Utóbbi „kritikára, szatírára és önreflexióra alkalmas formátum”, ami arra készítt, hogy reflektáljunk a „normáinkra, értékeinkre, ideológiáinkra és a létünkre”.²⁶ Ennek alesete a társadalomkritika megfogalmazása vagy a dezinformációnak való kitétség ábrázolása dokumentumfilmes eszközökkel. Magyarországon ennek legismertebb példája az Igazi Mao²⁷ című áldokumentumfilm volt, de más országokban is ismert, lásd például az Olajfalók²⁸ című cseh vízgátfilm. De a dokumentumfilmben

²⁴ Heather C. Hughes – Israel Waismel-Manor: The Macedonian Fake News Industry and the 2016 US Election. *PS: Political Science & Politics*, Vol. 54., No. 1. (2021) 19–23.
<https://doi.org/10.1017/S1049096520000992>

²⁵ Bajomi-Lázár i. m. 7.

²⁶ Jamriskó Tamás: *Mockumentary – Műfaj a valóság határán – Audiomanipulációs gyakorlatok*. In: Čakır-Somlyai (szerk.) i. m. 56.

²⁷ Az Igazi Mao (1995): <https://www.imdb.com/title/tt0109184/>

²⁸ Olajfalók (Ropáci – 1988): <https://www.imdb.com/title/tt0096010/>

látható dezinformáció ölthet további formákat is, mint amilyen például a fogyasztói társadalom kritikáját bemutató Cseh álom²⁹ című alkotás.

Társadalmi vonatkozásait illetően fontos motiváció lehet még a dezinformációra a szexuális vágykeltés vagy az erre alkalmas tartalmakkal való visszaélés. E mögött gyakorta szakítás miatti bosszút, gyerekes csínytevést vagy üzleti érdeket találni. Az interneten megosztott *deepfake* pornó komoly problémát jelent, mivel traumatizálhatja azokat, akiket az ilyen videókra vagy képekre montíroznak és tönkretelheti az életüket.³⁰ De komoly törést okozhat azok életében is, akiket ilyen tartalmak készítése miatt perbe fognak – a közelmúltban idehaza és világszerte is hallani ilyen esetekről iskolások körében.³¹

Ha megvizsgáljuk a hamisítások történetét, akkor számos további motivációt is láthatunk:³²

- a hatalom megszerzése vagy megtartása (például Constantinus állítólagos adománylevele a katolikus egyháznak)
- presztízs és tudományos elismertség (tudományos hamisítások, például a formosai – tajvani – útleírás George Psalmanazar tollából, vagy a piltdowni ember, amivel Charles Dawson írta be magát a tudománytörténetbe)
- életben maradás (pl. okmányok hamisítása a II. világháborúban)
- a titkosszolgálatok, katonai hírszerzés megtévesztő műveletei (háborús veszteségek eltitkolása és az ellenség manipulálása például a második világháború partraszállásának helyszínét illetően)

Tehát a motivációt illetően elég sokszínű képet látni. De vajon tud-e az MI maga dezinformációt létrehozni, tehát lehet-e saját motivációja mások megtévesztése? Az MI-kutatás elmúlt közel hetven évét domináló standard modell ezt a lehetőséget kizárja. E szerint ugyanis az MI „a helyes dolgot művelő ágensek tanulmányozására és építésére összpontosított. A helyes dolgot az ágensek általunk kiosztott célkitűzése adja meg.”³³ Vagyis az MI-nek nem lehet saját célkitűzése és így motivációja sem, azt tőlünk, emberektől kapja. Ugyanakkor időről-időre röppennek fel szenzációs hírek, amelyek ennek az ellenkezőjéről szólnak,³⁴ de az MI saját motivációja inkább kivételnek és bizonyításra váró tulajdonságnak tűnik, mintsem valóságnak.

²⁹ Cseh álom (Český sen – 2004): <https://www.imdb.com/title/tt0402906/>

³⁰ Karen Hao: Deepfake Porn is Ruining Women’s Lives. Now the Law May Finally Ban it. *MIT Technology Review*, 2021. február 12. <https://tinyurl.com/5n746f2h>

³¹ Ld. pl. RTL: Gyermekepornográfának is minősülhet az általános iskolai osztálytársról készült deepfake videó. *RTL*, 2024. május 16. <https://tinyurl.com/2xdrh9k2>; vagy Caroline Haskins: Florida Middle Schoolers Arrested for Allegedly Creating Deepfake Nudes of Classmates. *Wired*, 2024. március 8. <https://tinyurl.com/y8pcz95f>

³² A példák a brüsszeli Európai Történelem Háza utazó kiállításából származnak: *Valódi hamisítvány: A hamisítvány és a hamisítás története*. <https://tinyurl.com/4aheym2p>

³³ Stuart Russell – Peter Norvig: *Mesterséges intelligencia – Modern megközelítésben*. Budapest, Panem Könyvek, 2023. 5.

³⁴ Például Billy Perrigo: Exclusive: New Research Shows AI Strategically Lying. *Time*, 2024. december 18. <https://time.com/7202784/ai-research-strategic-lying>

Annak ellenére, hogy a motivációk sokrétűek a dezinformáció mögött, a cikk fókuszaltságának megőrzése érdekében és a terjedelmi korlátok miatt a továbbiakban elsősorban a politikai motivációval jellemezhető dezinformációs jelenségeket vizsgálom meg mélyebben. Ezek ugyanis azok, amelyek egyúttal nagyobb közönséghez is eljutnak és átfogó társadalmi hatással is bírnak, valamint a szakirodalom is jelentős részben ezekre szűkíti le a dezinformáció vizsgálatát.

2.2. Második lépés: létrehozás

A dezinformációs lánc második lépését, a dezinformáció létrehozását érdemes tágan értelmezni. Wardle tanulmányában az információs zavarok hétlépcsős skáláját különíti el, ami alapján kijelenthető, hogy számos módon lehetséges dezinformációt létrehozni.³⁵ Az első a szatíra vagy paródia, amely nem ártó szándékú, de a kontextusából kiemelve félreérthető lehet.³⁶ A második a hamis kapcsolat, amikor a cím és a tartalom nincs összhangban, csökkentve az ezt a gyakorlatot folytató médiába vetett bizalmat. A harmadik a félrevezető tartalom, ahol manipulált képek vagy adatok torzítják el a valóságot. A negyedik a hamis kontextus, amikor egy valódi tartalmat félrevezető módon használnak fel, például egy videójátékból vett képpel illusztrálnak egy valódi háborús hírt. Az ötödik az impoztor tartalom, amikor egy hiteles forrást hamisan tüntetnek fel, például más nevében kommunikálnak. A hatodik a manipulált tartalom, például amikor utólag eltüntetnek valakit egy képről. A hetedik pedig a hamis tartalom, ahol mindent a dezinformáció céljából hoznak létre. Mint látható Wardle lépcsője egyre aktívabb bevonódást feltételez, míg az elején egy félreértésből is lehet dezinformáció, addig az utolsó lépésnél már minden a cél érdekében kreált.

A létrehozás, mint dezinformációs lépés a generatív mesterséges intelligencia képességének rohamos javulásával kezdett egyre nagyobb figyelmet kapni. A generatív MI segítségével ugyanis létrehozhatók szövegek, képek, hangok, mozgóképek is, amiket akár élő telefon- vagy videóhívásban is felhasználhatnak. Létező megoldások közé tartozik a videón látható arcok kicserélése és a *lip syncing* (ajkak szinkronizálása) a kimondott tartalommal összhangban. A The Verge az álhírek lehetséges jövőjét illusztráló március 2018-ban olyan videót tett közzé, amelyben egy amerikai komikus, Jordan Peele beszél, de a néző Obamát látja és hallja helyette.³⁷ A technológia képessége azóta csak javult, 2024 februárjában egy hongkongi bank vezető alkalmazottja 25 millió dollárt utalt át csalóknak egy olyan online meeting után, ahol rajta kívül mindenkit – a cég pénzügyi igazgatóját és minden további kollégát – MI személyesített meg.³⁸ Az MI-vel

³⁵ Claire Wardle: Fake News. It's Complicated. *First Draft*, 2017. február 16. <https://firstdraftnews.org/articles/fake-news-complicated/>

³⁶ Fontos, hogy a szatíra vagy paródia nem tekinthető önmagában dezinformációnak, mivel nem félrevezetés céljából hozták létre. Félrevezető akkor lehet, ha a kontextusából kiemelik, például amikor azt valóságosnak próbálják feltüntetni.

³⁷ James Vincent: Watch Jordan Peele use AI to make Barack Obama deliver a PSA about fake news. *The Verge*, 2018. április 17. <https://tinyurl.com/2f7mb9b9>

³⁸ Heather Chen – Kathleen Magramo: Finance worker pays out \$25 million after video call with deepfake 'chief financial officer'. *CNN*, 2024. február 4. <https://tinyurl.com/yeyvxyvc>

készített hamis videók annyira hitelesek lehetnek, hogy akár magukat a videóra hamisított szereplőket is elbizonytalaníthatják egy pillanatra, ha találkoznak a felvétellel, hogy vajon tényleg mondtak-e olyat, amit a szájukba adtak.³⁹

Mivel ma már akár pár másodpercnyi hang és mozgókép is elég lehet valakiről a hamis megszemélyesítéshez, fontos, hogy az MI-modellek korlátozva legyenek, szabályozott módon és csak vízjelezve lehessen tartalmakat létrehozni velük. Ugyanakkor feltehető, hogy az általában korlátozottan felhasználható zárt MI-modellekkel szemben az *open source* MI modellek sokkal szabadabban használhatóak. Sőt, a kellő erőforrással bíró, dezinformációt rendszeresen végző szervezetek feltehetően saját céljaiknak megfelelően feltanított vagy módosított MI eszközökkel is rendelkezhetnek, amelyeket nem korlátoznak etikai elvek vagy szabályozásoknak való megfelelési kényszer.⁴⁰

Természetesen egy MI által generált, vagy bármilyen eszközzel létrehozott dezinformáció csak akkor éri el a hatását, ha nem, vagy nehezen detektálható. Ahogy az az áldokumentumfilmekből is látható, a műfaji sajátosságokhoz, kommunikációs stílushoz kell alkalmazkodni ahhoz, hogy valami hihető legyen. Bajomi-Lázár szerint a dezinformációban is kell lennie hihető elemnek, tehát „az álhírek és az igaz hírek szembeállítása hamis dichotómiát teremt, mert egy hír sosem teljesen igaz vagy teljesen hamis”.⁴¹ Vagyis a létrehozásának fontos eleme a hihetőség kontextusának megteremtése.

Az MI által létrehozható dezinformáció paradoxona viszont az, hogy bár a generatív MI megjelenésével drasztikusan megnőtt a dezinformációhoz használható eszközök száma és szinte bárki számára elérhetővé váltak, mégsem tapasztaltuk 2022 óta az ilyen jellegű dezinformáció megsokszorozódását:

„Eddig azonban nem láttuk a mesterséges intelligencia alapú dezinformáció várt áradatát. A 2024-es európai parlamenti választások is viszonylag nyugodtan zajlottak, nagyobb mesterséges intelligencia által generált dezinformációs kampányok nélkül, és úgy tűnik, hogy a korábbi választások is viszonylag immunisak voltak a mesterséges intelligencia által generált dezinformáció hatására.”⁴²

Vagyis a várakozások ellenére a generatív MI okozta korszakváltás elmaradni látszik – legalábbis, ami a létrehozást illeti. Ahhoz, hogy ennek okait megértsük, a dezinformációs lánc további lépéseit kell megvizsgáljunk – mint látni fogjuk ugyanis, nem

³⁹ Ezen a rövid videón például Joe Biden elnök arról beszél 2023-ban, hogy saját magáról látott egy *deepfake* videót és az egy pillanatra őt is megtévesztette, amíg rá nem jött, hogy ő sosem mondott ilyesmit: Biden Says He's Seen 'Mind-Blowing' AI Deep Fakes of Himself. *Bloomberg Quicktake*, 2023. október 31. <https://www.youtube.com/watch?v=CIDXoUoL9Qo>

⁴⁰ Noémi Bontridder – Yves Pouillet: The role of artificial intelligence in disinformation. *Data & Policy*, Vol. 3., e32 (2021) <https://doi.org/10.1017/dap.2021.20>

⁴¹ Bajomi-Lázár i. m. 6.

⁴² Judit Bayer: Disinformation and AI in the AV Sector. In: *AI and the audiovisual sector: navigating the current legal landscape*. European Audiovisual Observatory, Strasbourg, 2024. 81. <https://tinyurl.com/3stnzt88> (saját fordítás, P.R.)

a dezinformáció létrehozása jelenti a legnagyobb kihívást a dezinformációt folytatni kívánó szereplők számára.

2.3. Harmadik lépés: terjesztés

A dezinformáció terjesztése az egyik leginkább kutatott téma a dezinformációs lánc egyes lépései kapcsán. Nem véletlenül, ugyanis ez a kulcsa annak, hogy egyrészt detektálásra kerül-e a dezinformáció (erről részletesen lásd a következő részt), másrészt milyen lesz a befogadása és az hogyan hat vissza magára a terjesztésre, végül, hogy milyen hatást ér el a dezinformáció (ezekről is hosszabban írok később). Tehát a terjesztés, detektálás, befogadás és hatás szorosan egymásba fonódnak és csak az elemezhetőség miatt választottam külön őket. Összefonódásuk tetten érhető például a miszinformáció (félretájékoztatás) és a dezinformáció fentebb már taglalt különbségében is (a miszinformáció a dezinformáció akaratlan terjesztése).

Krekó szerint a dezinformációt terjeszthetik meggyőződéses összeesküvés hívők, akik hatalom és eszközök nélkül frusztráltan marginalizálódtak. De terjeszthetik anyagi érdekből is, mint láttuk a macedón fiatalok példáján. Végül a hatalomban levő populista vezetők hatalmuk megszilárdítása végett. Viszont, ami kiemelten fontos, hogy az elítelt és a tömegek együttműködnek a dezinformáció terjesztésében:

„Éppen ezért téves az a megközelítés, amely az álhírek »fogyasztóit« a meggyőzés folyamatának ártatlan áldozataiként látatja. Lehet, hogy a dezinformációt az elit gerjeszti, de az elítelt és a »nép« együttműködnek egymással a törzsi mítoszok előállításában és fenntartásában.”⁴³

Mint később látni fogjuk a dezinformáció kapcsán a közösségi média és az ahhoz köthető MI alapú algoritmusok jelentik a korszakalkotó lényegi újdonságot. Ami ugyan nem köthető az új generatív MI-hez, de a közösségi média algoritmusai miatt mégis MI-központú jelenség. Ináncsi szerint a közösségimédia-tartalmak terjedésében a kulcs a moderáció (erről a detektálásról szóló következő részben írok bővebben is), ugyanakkor:

„[a] használt moderációk (egyéni szinten történő felülvizsgálatok, algoritmus alapú moderáció, vagy teljes moderálásmentesség, illetve az alternatív moderációk) nem alkalmasak arra, hogy megfelelően szűrjék a felhasználók által megosztott tartalmakat.”⁴⁴

Ennek oka több, egymással rokon jelenség, amelyeket *astroturfingnek*, szimulált tömegnek, összehangolt nem hiteles viselkedésnek és részvételi propagandának nevezünk. „Az astroturfing a szakirodalomban az online visszaélések egy altípusaként

⁴³ Krekó i. m. 296–297.

⁴⁴ Ináncsi Máttyás: Közösségi média és dezinformáció. *Military and Intelligence Cybersecurity Research Paper*, 2023/7. 16. <https://tinyurl.com/bdfepxc3>

jelenik meg, amikor politikai kampányok grassroots, spontán viselkedésnek tüntetik fel azt, amit a valóságban egy ember vagy egy szervezet hajt végre.⁴⁵ Tehát egyetlen, dezinformációt folytató személy vagy szervezet több hamis közösségimédia-profil is kezelhet, amivel egyrészt elfedi a dezinformáció koordinációját, másrészt miközben spontánnak tűnő tömeget hamisít a tevékenység mögé, a meglévőnél jóval nagyobb társadalmi relevanciát is sugall. Bár a jelenség a közösségi média előtt is létezett (például fizetett tüntetők), de a közösségi média részvételi jellege miatt sokkal könnyebbé tette a dezinformációs kampányok menedzselését. Ezt a tevékenységet a közösségimédia-platformokat üzemeltető cégek összehangolt, nem hiteles viselkedésnek tekintik – például, amikor fiókok százalai vagy ezrei osztják meg közel egyidőben ugyanazt az üzenetet.

Az igazi problémát azonban az jelenti, hogy a közösségi médiában a dezinformációt nem csak a létrehozók terjesztik koordináltan, hanem ebbe bárki – akár végtelenül is – bekapcsolódhat, aminek következménye, hogy „a közösségi média révén a »propagandista« és a »célközönség« közötti határ elmosódik, a korábban passzív közönség pedig egyre jelentősebb szerepet játszik a propagandisztikus tartalmak terjesztésében.”⁴⁶ Ezt nevezi a szakirodalom részvételi propagandának.⁴⁷ Ugyanakkor az a jelenség évtizedek óta ismert, hogy a politikai vagy marketing kommunikációban a létrehozó a célcsoport tagjainak erőforrásait használja az üzenetek terjesztésére. Csakhogy ez a közösségi média előtt legfeljebb e-mailt és SMS-t jelentett és csak tudatos mozgósítás árán tudott igazán tömeges méreteket öltetni.⁴⁸ Ami új, hogy „az üzenetek továbbítását programozott botok és valós emberek együttesen végzik el. Ez utóbbi pedig azzal jár, hogy lehetetlenné válik szétválasztani, hol végződik a hamisítás (koordináció) és hol kezdődik a hitelesség (organikus viselkedés).”⁴⁹ Ezzel a hagyományos tömegmédián túl létrejött egy olyan új tér, ahol a politikai (és gazdasági) üzenetek hatékonyan eljuttathatók és terjeszthetők a közönség aktív bevonásával. Mindez persze kétélű:

„[a] részvételi kultúra így egy lehetőség, de egy kitétség is a politika számára: minél többen felkapják, annál hitelesebbé válik az információ, a megfelelő üzenetek kiválasztása és továbbítása felett azonban az intézményes szereplők könnyen elveszíthetik a kontrollt.”⁵⁰

Ennek a hatékonyságát két további dolog is jelzi. Egyrészt az így terjesztett dezinformációs tartalmak akár a Google keresési lista élére is tudnak kerülni (kikerülve

⁴⁵ Német Szilvi: Szűrkezónás médiumok, álcivil kampányok és részvételi propaganda a közösségi média platformokon. In: Çakır-Somlyai (szerk.) i. m. 68–69. [A továbbiakban: Német (2023a).]

⁴⁶ Uo. 73.

⁴⁷ Alicia Wanless – Michael Berk: Participatory Propaganda – The Engagement of Audiences in the Spread of Persuasive Communications. *Social Media & Social Order, Culture Conflict 2.0 Conference*, Oslo. 2017 november. <https://tinyurl.com/ycxbyyh8>

⁴⁸ Ld. pl. Sükösd Miklós – Dányi Endre: E-mail és SMS a 2002-es választási kampányban. *eDemokrácia Műhely*, év nélkül. <https://www.edemokracia.hu/program/smskampany.php>

⁴⁹ Német (2023a) i. m. 73.

⁵⁰ Uo. 74.

a pusztán közösségi médiából) és onnantól kezdve hitelesíti őket a Google, mint az internetes keresés legfontosabb nyugati szereplője.⁵¹ Másrészt ezek a dezinformációs tartalmak „gyorsabban, messzebbre és szélesebb körben” terjednek, mint a cáfolatuk vagy az igazi hírek.⁵² Ezeket ugyanis nagyobb eséllyel és hamarabb osztják meg az emberek, mivel „a hamis hírek gyakran érdekesebbek, fontosabbnak tűnnek, rájátszanak az előítéleteinkre, azaz sokkal inkább megosztásra sarkallnak, mint az igazság.”⁵³

A szűrőbuborék (*filter bubble*) és a visszhangkamra (*echo chamber*) fogalmak is segíthetnek megérteni a közösségi médiában terjedő részvételi propaganda sikerességének okait, bár évtizedes vita folyik mindkét jelenségről.⁵⁴ Röviden, a közösségi média személyre szabott módon, az érdeklődésünk és korábbi aktivitásaink mentén szűri a nekünk szánt információkat (szűrőbuborék). Ezen felül pedig olyan személyekkel vesz körbe minket, akik a saját magunk által is hangoztatott, hozzánk hasonló véleményeket osztják (visszhangkamra). Ugyan a közösségi média használata a tájékozódásban nem feltétlenül kizárólagos, de a két jelenség abban mégis meghatározó, hogy mit láthatunk magában a közösségi médiában, így közvetve arra is hatással bír, hogy mivel keveredünk interakcióba, például mit lájkolunk vagy osztunk meg és így terjesztünk tovább. Mint később látni fogjuk, emiatt a befogadásra és a társadalmi következményekre is érdemi hatásuk lehet.

A botrányaiba belebukott Cambridge Analytica gyakorlata arra mutatott rá, hogy a dezinformációk terjesztésének hatékonyságát növeli a személyre szabás, avagy a mikrocélzás (*micro targeting*) is.⁵⁵ A felhasználók aktivitásainak követésével a közösségimédia-plafomok lehetővé teszik hirdetőik és tartalomkészítőik számára, hogy meghatározott demográfiai vagy érdeklődési csoportoknak lehessen üzeneteket eljuttatni, így a dezinformáció könnyebben találja meg az arra fogékony közönséget. Az ilyen magas fokú célzás nagyobb manipulációt képes elérni és megsokszorozza a megtevesztő tartalmak hatását.⁵⁶ Nem véletlen, hogy a Cambridge Analytica által használt dezinformációs technikákat az információs hadviselés eszközei közé sorolhatjuk.⁵⁷

Természetesen a közösségi médiának a dezinformáció terjesztésében való felértékelődésével nemcsak a kutatók vannak tisztában, hanem azok is, akik dezinformációs kampányokat folytatnak. Például Oroszország több közösségimédia-plafomon is aktív és szisztematikus módon terjeszti a dezinformációt álhírgyárak segítségével. Ennek

⁵¹ David Gilbert: How Disinformation From a Russian AI Spam Farm Ended up on Top of Google Search Results. *Wired*, 2024. július 9. <https://tinyurl.com/5n86ptb7>

⁵² Pavel Nikolic: A visegrádi négyek és a dezinformáció. *CEU Democracy Institute*, év nélkül. <https://cmds.ceu.edu/visegradi-negyek-es-dezinformacio>

⁵³ Uo.

⁵⁴ Erről kiváló cikk olvasható Gálik Mihály tollából: Gálik Mihály: A hálózati hírmédia sajátosságai, különös tekintettel a visszhangkamra és a szűrőbuborék-jelenségre. *In Medias Res*, 2019/2. 330–342. <http://real.mtak.hu/108287/1/imr-2019-02-08-jav-1.pdf>

⁵⁵ Jessica Dawson: Microtargeting as information warfare. *The Cyber Defense Review*, Vol. 6., No. 1. (2021) 63–80. <https://www.jstor.org/stable/26994113>

⁵⁶ Bontridder–Poulet i. m.

⁵⁷ Dawson i. m.

egyik központja a szentpétervári Internet Research Agency,⁵⁸ ami 2013 óta működik és kiterjedt hálózatot tart fenn fizetett propagandisták (mai kifejezéssel trollok) és hamis közösségi fiókok alatt rejtőző bothadseregek segítségével.⁵⁹ De az amerikai titkosszolgálat szerint a döntően meghatározó orosz tevékenységen túl Kína és Irán is folytatott dezinformációs tevékenységet például a legutóbbi elnökválasztás idején, hogy aláássa az amerikai demokratikus intézmények működését és társadalmi megosztottságot szítson.⁶⁰

2.4. Negyedik lépés: (detektálás)

A dezinformációs lánc negyedik lépése a detektálás, ami annyiban tér el a többi lépéstől, hogy nem mindig történik meg (emiatt szerepel a fejezet címében is zárójelben). Mindez érthető is, hiszen a létrehozó célja a megtévesztés, amit könnyebben ér el, ha nem azonosítja dezinformációként az üzenetét. Fontos, hogy a detektálás csak a terjesztéssel és a befogadással szoros összefüggésben értelmezhető, ezt a lépést csak az elemezhetőség miatt választottam külön ezektől.

A detektálás több szinten zajlik, mikroszinten maguk az egyes befogadók is érzékelhetnek valamit dezinformációnak az olyan gyanús jelek alapján, mint a szerző, a dátum vagy a hivatkozások hiánya, esetleg a korábban már bizonyítottan dezinformációt terjesztő források szerepeltetése.⁶¹ Ugyanakkor egyéni szinten rendkívül nehéz az összehangolt nem hiteles viselkedés detektálása, mivel jellemzően egyetlen profillal látjuk ezeket a tartalmakat és nem látjuk a terjedési mintázatokat, a részt vevő hálózat működését – erre később még visszatérek. Az egyén így a dezinformációt elsősorban annak tartalma felől hiteltelenítheti. Szerencsére ebben nincs egyedül, 2023-ban közel 400 független tényellenőrző szolgáltatás volt elérhető, de a fogyasztók megfelelő médiatudatossága és eszközhasználatára nélkül ezek hatékonysága alacsony maradhat.⁶² Mivel a részvételi propaganda megerősödése miatt a dezinformáció terjesztése rendszerint a hagyományos tömegmédiára megkerülésével a közösségi médiában zajlik, felértékelődik a külső tényellenőrzés szerepe (azaz, ami nem szerkesztőségekhez köthető, amelyek a publikálás előtt jó esetben leellenőrzik egy információ hitelességét).

A tényellenőrzés kihívása három rétegű, egyrészt kevés a tényellenőrzéshez szükséges szervezet és emberi kapacitás, szemben az interneten áramló hatalmas mennyiségű (dez)információval szemben. Másrészt a tényellenőrzés a közösségi médiában nem előzi meg a terjesztést, mint a hagyományos média esetén, hanem versenyt fut

⁵⁸ Az IRA tevékenységéről ld. még: Philip N. Howard – Bharath Ganesh – Dimitra Liotsiou – John Kelly – Camille François: *The IRA, Social Media and Political Polarization in the United States, 2012-2018*. Computational Propaganda Research Project. University of Oxford, 2018. <https://tinyurl.com/3xmf2nw>

⁵⁹ Çakır-Somlyai Özge: Dezinformáció és demokrácia – A jelenlegi helyzet és a jövőbeli lehetőségek vizsgálata. In: Çakır-Somlyai (szerk.) 109.

⁶⁰ *100 Days Until Election 2024*. Office of the Director of National Intelligence, 2024. július 29. <https://tinyurl.com/2rakp347>

⁶¹ Veszelszki Ágnes: Az álhírek extra- és intralingvális jellemzői. *Századvég*, 2017/2. 51–82. <https://tinyurl.com/3b3ka943>

⁶² Horváth Kata: Álhírek leleplezése – Tényellenőrzés a világban és Magyarországon. In: Çakır-Somlyai (szerk.) i. m. 77.

azzal és mint fentebb láthattuk, a dezinformáció érdekessége folytán gyorsabban és messzebbre terjed. Harmadrészt a tényellenőrzés maga is a terjesztési folyamat részévé válik, a tényellenőrzés eredményének megítélése a dezinformációt akaratlanul vagy tudatosan terjesztő felhasználóktól függ, vagy elfogadják azt vagy sem. Ezért aztán nem elég az objektív tényellenőrzés eredménye ahhoz, hogy egy dezinformáció 'lebukjon' és onnantól kezdve mintegy varázsütésre ne terjedjen tovább. Német kutatása alapján például a tényellenőrzés műfajának megítélésében a politikai szempontok domináltak a Lakmusz tényellenőrző Facebook oldalának esetében: „a fősodró média napirendjébe illeszkedve a tényellenőrző cikkek kevesebb bírálatot váltottak ki az olvasók körében, míg a politikai táborok közötti polarizációt növelhették.”⁶³

Ugyanakkor nem mindenki szerint jó ötlet tényellenőrzéssel, vagy a hagyományos médiában bevált emberi moderálással küzdeni a dezinformáció ellen. Az X (korábban Twitter) és a Facebook 2024-2025-ben szakított a korábbi gyakorlatával és már nem alkalmaz tényellenőröket, ehelyett egy úgynevezett *community notes* rendszerre álltak át. Ennek lényege a folkszonómia, vagyis a felhasználók bevonása a tartalmak megítélésébe és az általuk jelentett tartalmak vizsgálata. Ennek hatékonysága azonban megkérdőjelezhető.⁶⁴ A Facebook esetében egyelőre nem lehet tudni, hogy vajon a 2025 januárjában hivatalba lépő Trump elnökkel való jó kapcsolat kialakítása áll az áttérés hátterében, vagy a Facebook tényleg tarthatatlannak tekintette a dezinformáció terjedésének tényellenőrzés alapú kontrollálását.

Ez már csak azért sem egyértelmű, mert van, aki már évek óta a tényellenőrzés hiányosságaira figyelmeztet. Tim O'Reilly, a webkettő fogalom atyja szerint a közösségi médiában terjedő csalásokat emberi eszközökkel nem lehet kivédeni a csalások kiterjedtsége és gyorsasága miatt. Ennek oka, hogy csak egy gép képes belépni egy másik gép OODA (*Observe – Orient – Decide – Act* / Megfigyel – Elemez – Dönt – Cselekszik) ciklusába és megzavarni azt, egy ember nem elég gyors hozzá. O'Reilly szerint azonban az algoritmusok mégis több mindent megtehetnének a hírek kapcsán, például:

- ahogy a részvényeknél számolnak fundamentális értéket, ami a tényleges érték mögött van, úgy lehetne a híreknél is számolni ilyesmit,
- megkereshetnék a hírek eredeti forrásait, hogy nem mondanak-e ellent a megosztott hír tartalmának,
- megnézhetnék, hogy mennyien hivatkoznak rá, hitelesnek tekinthető-e a közlő oldal,
- leellenőrizhetnék, hogy statisztikailag releváns módon tálalják-e az információkat (például a bűnügyi statisztikákat minden esetben ugyanannyi lakosra vetítve közlik-e),

⁶³ Német Szilvi: „Értelmezni se tudom, mi az a liberális tényellenőr” – Facebook-hírkomentárok a tényellenőrző újságírásról. *Médiakutató*, 2023/3. 79. <https://tinyurl.com/y7bnpwzy> [A továbbiakban: Német (2023b).]

⁶⁴ Yuwei Chuai – Haoye Tian – Nicolas Pröllochs – Gabriele Lenzini: Did the Roll-Out of Community Notes Reduce Engagement With Misinformation on X/Twitter? *Proceedings of the ACM on Human-Computer Interaction*, Vol. 8., No. CSCW2. (2024) 1–52. <https://doi.org/10.1145/3686967>

- megíthetnék, hogy kattintásvadász-e a címük, azaz olyan, ami nem tükrözi rendesen a tartalmat,
- ellenőrizhetnék, hogy több, független forrás is lehozta-e már a hírt,
- megkérdezhetnék a felhasználókat, hogy biztos tovább akarnak osztani egy posztot anélkül, hogy rákattintottak volna a benne szereplő linkre,
- végül összevethetnék a cikk tartalmát az eredeti forrással, például, ha törvényre hivatkozik, akkor a törvény szövegével.⁶⁵

Persze ezeket az emberek is megtehetnék. De O'Reilly szerint az emberek ezeket a gyakorlatokat eddig sem követték és nem is fogják megváltoztatni a hírolvasási szokásaikat csak úgy. Ráadásul, ha valami olyat olvasnak, amit el akarnak hinni, akkor pláne nem fognak energiát fektetni abba, hogy megkeressék annak esetleges cáfolatát.

És ezzel el is érkeztünk a detektálás kapcsán az egyik legfontosabb, ha nem a legfontosabb kérdéshez: tényleg lehet MI-vel küzdeni az MI által támogatott dezinformáció ellen? Az utóbbi évtizedben szinte MI 'fegyverkezési verseny' bontakozott ki a dezinformáció terjesztésében és detektálásában, illetve korlátozásában.⁶⁶ Sedova és munkatársai a RICHDATA piramis leírásával visszafejtették a közösségi média alapú politikai dezinformációs kampányok működését, miközben akaratlanul egyfajta 'receptet' is adtak ahhoz, eszközöket is nyújtottak az ellene való küzdelemhez. A piramis leírja a dezinformációs kampányok létrehozását, felépítését, terjesztését és a közönség bevonását:

„A dezinformációs kampányokat több szakaszra bontjuk. A felderítés [Reconnaissance] során az operátorok felméri a környezetet, és megértik a közönséget, amelyet manipulálni próbálnak. Infrastruktúrára [Infrastructure] van szükségük – hírvivőkre, hihető személyiségekre, közösségi média fiókokra és csoportokra –, hogy a narratíváikat továbbítsák. Az üzeneteik magva, gyökerezése és növekedése érdekében elengedhetetlen a tartalom [Content Creation and Hijacking] szüntelen áramlása, a posztoktól és hosszú olvasmányoktól kezdve a fotókig, mémekig és videóig. Az internet áramlatában való megjelenés [Deployment] után a dezinformáció ezen egységeit botok, platform-algoritmusok és social-engineering technikák erősítik fel [Amplification], hogy a kampány narratíváit terjesszék. A dezinformáció szétbontása azonban nem mindig elegendő: a széles körű hatás a tudatlan felhasználókkal való tartós kapcsolattartáson keresztül, trollkodással [Troll Patrol] érhető el – ez a dezinformációs közelharc megfelelője. A végső szakaszban a dezinformációs művelet azáltal válik aktuálissá [Actualization], hogy megváltoztatja a tudat-

⁶⁵ Tim O'Reilly: *WTF – Miért rajtunk múlik, hogy mit hoz a jövő?* Budapest, Typotex, 2018.

⁶⁶ Fátima C. Carrilho Santos: Artificial Intelligence in Automated Detection of Disinformation: A Thematic Analysis. *Journalism and Media*, Vol. 4., No. 2. (2023) 679–687. <https://doi.org/10.3390/journalmedia4020043>

lan célpontok gondolkodását, vagy akár cselekvésre mozgósítja őket, hogy károszt szítsanak.”⁶⁷

A RICHDATA annak belátásában is segít, hogy több ponton lehetne beavatkozni MI segítségével, ezek közül az első pont kétségkívül a dezinformáció detektálása lenne. Bayer szerint több fajta módon lehet az MI-t a dezinformáció detektálására használni: a terjesztők azonosításában, a terjedési mintázat detektálásában és a tartalom felismerésében.⁶⁸ Ezeket szerinte érdemes vegyíteni, például lehetséges detektálni a botokat, amelyek közel egyszerre osztják meg ugyanazt az üzenetet, gyanús az is, ha egy közösségimédia-tartalomnak több megosztása, mint kedvelése van, mert rendszerint fordítva szokott lenni. Viszont azzal, hogy a botok keverednek a valós emberekkel, sokszor nagyobb követői táborokkal rendelkező influenszerekkel és celebritásokkal (akik akaratlanul vagy meggyőződésből terjeszthetik a dezinformációt), meglehetősen nehézkessé válik az MI használata ezeknek a hibrid dezinformációs hálózatoknak az automatizált kezelésében (például, hogy mely profilokat *shadow banning*olják⁶⁹, függesszék fel vagy töröljék véglegesen). Ezen túlmenően Bayer szerint egyelőre az MI még nem elég fejlett ahhoz, hogy a tartalom elemzése alapján azonosítsa be a dezinformációt. Ehhez a számtalan beszélt emberi nyelv finomságait (például ironia), kulturális kontextusait és politikai vonatkozásait is ismernie, és ezt a tudást magabiztosan alkalmaznia is kellene. A legnagyobb kihívás azonban az, hogy ezeknek a mesterséges intelligencia által támogatott módszereknek a kombinálása sem elég hatékony, amikor azonnali beavatkozásra lenne szükség a futótűzszerűen terjedő dezinformációval szemben, például „amikor a dezinformáció a tőzsdei folyamatokat kívánja manipulálni, háború idején vagy közvetlenül a választások előtt.”⁷⁰ Egyelőre a gyors és megbízható MI-alapú dezinformációdetektáló rendszerek kifejlesztése még várta magára, bár a nagy nyelvi modellekkel már elkezdődtek a tesztek.⁷¹

De felmerül a kérdés, hogy egyáltalán szükségünk van-e a detektálásra, változik-e tőle valami, érdeklie az embereket, ha kiderül egy információról, hogy hamis? O’Reilly szerint a történelem, a politika vagy az üzlet nem fekete-fehér, nem igaz vagy hamis, hanem olyan mértékben igazak vagy hamisak, amennyire annak hisszük azokat – ezt nevezi ő meggyőződésnek.⁷² Az emberek pedig nem feltétlenül azért hisznek el valamit, mert az igaz, hanem mert el akarják hinni. Ezzel pedig elérkezünk a befogadás kérdéséhez.

⁶⁷ Katerina Sedova – Christine McNeill – Aurora Johnson – Aditi Joshi – Ido Wulkan: *AI and the Future of Disinformation Campaigns Part 1: The RICHDATA Framework*. Center for Security and Emerging Technology, December 2021. 1. <https://tinyurl.com/5n78ucm2> (saját fordítás P.R.)

⁶⁸ Judit Bayer: AI and disinformation – The technological and legal limits of content regulation in the context of disinformation. *Legal Tech*, 2022/2. 1–8. <https://tinyurl.com/34pphen>

⁶⁹ A *shadow banning* (magyarul talán *árnyék némitásnak* lehetne fordítani) lényege, hogy továbbra is lehet használni a fiókot, de a közzétett tartalmakat csak az azokat közzétevő felhasználója látja, más nem, és erről az érintett nem is feltétlenül szerez tudomást.

⁷⁰ Bayer (2022) i. m. 3.

⁷¹ Bayer (2024) i. m. 89.

⁷² O’Reilly i. m. 261.

2.5. Ötödik lépés: befogadás

A dezinformációs lánc ötödik lépésénél, a befogadás esetében a kiindulópontot az emberi észlelés és információfeldolgozás megértése adja. Ahogy Krekó kiemeli az evolúció során az szolgálta a túlélésünket, ha a saját érzékszerveinkre hagyatkoztunk és alapvetően gyanakvóak voltunk. Ez azonban elvezetett ahhoz, hogy mára viszonylag könnyen manipulálhatóak lettünk, főleg mivel „a legtöbb információt közvetlenül szerezük, és a szövegeket, a képeket, illetve a hangot egyaránt könnyebb manipulálni, mint korábban bármikor.”⁷³

Persze ettől még kell, hogy a dezinformációnak legyen hihető része, mint azt korábban a létrehozás lépésénél már említettük:

„az álhírekben is van némi igazság: egy teljesen légből kapott hírt a befogadók nem hinnének el. A Central Park állatkertjéből sosem szöktek meg a vadállatok, de valóban van ott állatkert, Cion bölcséinek sosem volt titkos jegyzőkönyvük, de cionista világkongresszust valóban tartottak... és így tovább. A koholmányokat az ismert elem teszi hihetővé, mert a befogadó így az új információt be tudja illeszteni a világról alkotott képébe.”⁷⁴

De a dezinformáció esetében a hihetőségen túl legalább ennyire kulcsfontosságú, hogy mit akarunk és kinek elhinni. Miért hisszük el a tévhiteket és ragaszkodunk hozzájuk makacs módon? Hammer szerint⁷⁵ Haidt moráleméleti kategorizációja mentén lehet megérteni a jelenséget:⁷⁶

1. az emberek úgy érzik joguk van eldönteni, hogy kinek és mit higgyenek el;
2. az emberek számba veszik a kockázatokat és veszélyeket, amik az általuk fontosnak ítélt személyeket (pl. család) és dolgokat fenyegetik;
3. vannak dolgok, amikről az emberek nem hajlandók lemondani vagy veszélynek kiténni őket (pl. a vallásuk vagy természeti környezetük);
4. fontos, hogy mit gondolnak az emberek arról, mások mennyire bánnak méltányosan velük;
5. miről vagy kiről érzik, hogy hűséggel tartoznak neki (pl. lakóhelyük közössége, football csapatuk stb.);
6. végül meghatározó az is, hogy milyen tekintélyt ismernek el.

Vagyis a befogadás kontextusa nélkül nem lehet megérteni magát a dezinformációt. Azok a dezinformációk hullanak termékeny talajra, amelyek feljogosítják az embereket,

⁷³ Krekó i. m. 267–268.

⁷⁴ Bajomi-Lázár i. m. 8.

⁷⁵ Hammer Ferenc: Osztálytermi fórum a dezinformációról. In. Çakır-Somlyai Özge (szerk.): *A DEZINFORMÁCIÓ JELENSÉGE – tanulmánygyűjtemény*. Budapest, Mérték Füzetek 33. szám, 2023. 94–95. https://mertek.eu/wp-content/uploads/2023/12/Mertek-fuzetek_33_1.pdf

⁷⁶ Jonathan Haidt: *The righteous mind: Why good people are divided by politics and religion*. Vintage, 2012.

hogy higgyenek bennük, az emberek számára fontos személyekkel és dolgokkal kapcsolatos veszélyekről szólnak, (látszólagos) méltánytalanságra hívják fel a figyelmet, olyanoktól erednek, akik iránt hűségesek a dezinformációval megcélzott személyek és általuk elfogadott tekintély is hitelesíti a tartalmukat. Krekó hasonló következtetésre jut összeesküvésekről és dezinformációról szóló könyvében, amikor a következőt írja: az „hogy mit tekint valaki összeesküvés-elméletnek vagy álhírnek, nagyban függ már előzetesen meglévő véleményétől, illetve attól, hogy milyen tekintélyeket fogad el hiteles forrásnak.”⁷⁷

Mindez azt is megmagyarázza, hogy a közösségi média algoritmusok működésén túl miért lehetnek annyira sikeresek a visszhangkamrák és a szűrőbuborékok. Eltérő módon reagálunk ugyanis ugyanazokra az információkra attól függően, hogy mi fontos nekünk, emiatt mindenki (legalább egy kicsit) a saját valóságában él, ami már a közösségi média megjelenése előtt is jellemző volt ránk, emberekre.

Emiatt felmerül az a kérdés is, hogy szükséges-e egyáltalán a médiának (vagy bármelyik külső tényellenőrző szervezetnek) *deepfake* ellenőrző kapacitást kiépítenie. És ez nem pusztán azért merül fel, mert ezek egyelőre nem működnek megbízható módon: az egyik szoftver azt mondja valamiről, hogy *deepfake* hamisítvány, a másik pedig azt, hogy nem az.⁷⁸ Hanem mert emiatt a bizonytalanság miatt a végén az olvasónak vagy a nézőnek magának kell úgyis eldöntenie, hogy kinek higgyen. Az emberek pedig egyre inkább abban hisznek, amiben akarnak, kiválogatják azokat az információkat, amik megerősítik a véleményüket, hitüket, ideológiai beállítódásukat. Ahogy Krekó írja: az összeesküvés elméletekben hívők befogadása megváltozik. Nem azért hiszik el valamiről, hogy igaz, mert az annyira tökéletesen manipulált *deepfake* lenne, hanem mert el akarják hinni.⁷⁹

De a dezinformáció személyes befogadásában legalább ennyire fontos a közösségi média működési mechanizmusa és üzleti modellje, illetve annak következményei a befogadás körülményeire. Ináncsi szerint a közösségi média célja az erős érzelmi reakciók kiváltása, mivel ez növeli a felhasználói interakciókat. A viták elősegítik a tartalomtermelést, ami a platformok működésének alapja és a dezinformáció terjedését fokozza, hogy az álhíreket megosztó felhasználók erősebb érzelmi reakciókat mutatnak. Mindez kihívást jelent a moderáció számára is, mivel, ha a közösségimédia-platformok az interakciót vitákra építik, az kedvezhet a félrevezető tartalmak elterjedésének.⁸⁰ Tehát a közösségimédia-oldalak üzleti modelljük miatt (ami az ingyenességen és a reklámbevételen alapul) abban érdekeltek, hogy minél hosszabb ideig fenntartsák a figyelmet, mert azt értékesítik a reklámozók felé. Ezért algoritmusaik annak adnak

⁷⁷ Krekó i. m. 261.

⁷⁸ Laura Stroebel – Mark Llewellyn – Tricia Hartley – Tsui Shan Ip – Mohiuddin Ahmed: A systematic literature review on the effectiveness of deepfake detection techniques. *Journal of Cyber Security Technology*, Vol. 7., No. 2. (2023) 83–113.
<https://www.tandfonline.com/doi/full/10.1080/23742917.2023.2192888>

⁷⁹ Krekó i. m. 26.

⁸⁰ Ináncsi i. m. 14–15.

teret, ami ott tartja a felhasználókat, ezek pedig a megosztó, érzelmeket, bevonódást és vitát kiváltó tartalmak, köztük maga a dezinformáció is.

A korábbiak során rendre külön megvizsgáltam az MI szerepét a dezinformációs lánc adott lépésében, illetve az MI által támogatott lehetséges beavatkozási pontokat (bár a dezinformáció elleni küzdelemről később részletesen is írok). A dezinformáció percepciója, feldolgozása és befogadása kapcsán azonban a már fentebb említetteken túl nincs kiemelt szerepe az MI-nek, mivel a befogadás teljes egészében emberi folyamat. Az MI sokkal inkább meghatározó tud lenni a befogadással szorosan összefonódó terjesztés és detektálás lépéseiben, mint magában a befogadásban.⁸¹

2.6. Hatodik lépés: hatás és (nem szándékolt) következmények

„Az álhírek és összeesküvés-elméletek nem elsősorban önmaguk miatt érdekesek, hanem az általuk kiváltott hatás miatt.”⁸² A dezinformációs lánc működésének vizsgálata tehát fontos ugyan, de pusztán azt a célt szolgálja, hogy közelebb férhessünk a társadalmi hatások mélyebb megértéséhez.

A dezinformációnak több fajta módon is kategorizálhatjuk a hatását, például:

- van tétje (pl. háborús uszítás) vagy nincs igazi tétje (pl. a pufi dzsekis pápa képe⁸³);
- mikroszintű: személyes – mezoszintű: egyes intézményekre korlátozódó – makroszintű: társadalmi – metasztintű: például annak vizsgálata, hogy okoz-e korszakváltást;
- lokális, regionális, országos vagy nemzetközi hatású;
- illegális vagy nem illegális;
- milyen mértékben okoz kárt: a teljesen ártalmatlantól a globális szintű károkozásig;
- hol, milyen területen fejt ki a hatását (pl. csoportközi erőszak, terrorizmus, demokratikus folyamatok, egészségügy, környezetvédelem, politikai szélsőségek).⁸⁴

A cikk terjedelmi korlátai miatt az alábbiakban a nagy téttel bíró, társadalmi vagy metasztintű változásokat okozó, országos vagy nemzetközi hatású, illegális és globális károkozásra képes dezinformációs tevékenységekre koncentrálok.

Ezek alapján kijelenthető, hogy a dezinformáció „drasztikus következményekkel is járhat, melyre példa amikor az Egyesült Államokban a Capitol(sic!) ostromához és

⁸¹ Az egy érdekes kérdés, hogy ez a jövőben vajon megváltozhat-e. Ray Kurzweil 2024-ben megjelent könyvében (*Singularity is nearer – When we merge with AI*. Vintage, 2024.) a 2040-es évek közepére teszi az ember(i agy) és a gép(i gondolkodás) egyesülését. Ha ez megvalósulna, akkor ez kétségtelenül megváltoztatná az információk befogadásának folyamatát is.

⁸² Krekó i. m. 105.

⁸³ Gergely Máté: Pufi dzsekis Ferenc pápa és társai: így leplezzük le az MI-képeket. *Világgazdaság*, 2024. április 13. <https://tinyurl.com/3r4hry7a>

⁸⁴ Ezekről bővebben ld. Krekó i. m.

megszállásához vezetett az álhír terjedés.”⁸⁵ Bajomi-Lázár szerint „[a] társasmedia-platformokon terjesztett hamis információk szerepet játszhattak a 2016-os amerikai elnökválasztás és a brexit végeredményében is.”⁸⁶ A dezinformáció képes lehet aláásni a demokratikus intézményeket és magát a demokráciát is, „[a] 2016-os események, köztük a Brexit-népszavazás és az amerikai elnökválasztás aláhúzták a dezinformáció veszélyét és a demokratikus társadalmak működését megzavaró potenciálját.”⁸⁷

Bayer szerint a dezinformációs kampányok egyre szélsőségesebbé váltak az utóbbi években.⁸⁸ 2015–2016-ban az amerikai választások és a Brexit-népszavazás befolyásolásával kezdődött, majd a COVID-19-cel kapcsolatos álhírek sokak egészségügyi döntéseit eltorzították. A 2021. január 6-i Capitolium elleni támadás során öt ember meghalt, míg az Oroszország, Ukrajna és a Nyugat közötti feszültségeket tápláló dezinformáció háborúhoz vezetett.

Barts szerint az internet és a közösségi média alapjaiban változtatta meg az információ terjesztését, lehetővé téve a dezinformáció gyors és szervezett terjedését.⁸⁹ Információs túlterheltség esetén az emberek inkább a saját csoportjuk tagjaira támaszkodnak, nem pedig szakértőkre.⁹⁰ Fontos, hogy a dezinformáció ismétlése ismerősségérzetet kelt, ami idővel növeli az elfogadás esélyét (ezt nevezzük ‘illuzórikus igazsághatásnak’). Emellett az információ és forrása gyakran különválnak az emlékezetünkben, így a hamis adatok akkor is befolyásolhatják az attitűdjeinket, ha később kiderül róluk, hogy dezinformáción alapultak.

2016-ban Trump első megválasztásától nem függetlenül az Oxford Dictionaries az év szavának választotta a *post-truth* kifejezést. A fogalom ugyan nem 2016-ban született, hanem feltehetően 1992-ben, de az elmúlt évtized második felétől datálható széles körben való elterjedése, miután egyik évről a másikra hűsszorosára nőtt a használata. Jelentése szerint „[o]lyan körülményekre vonatkozó melléknév, amelyekben az objektív tények kevésbé befolyásolják a közvéleményt, mint az érzelmi vonzalmak.”⁹¹ A *post-truth*ra jellemző a dezinformáció és az összeesküvés elméletek széleskörű térnyerése. Ezek együtt elvezettek ahhoz, hogy még megosztottabbak legyenek a társadalmak és ugyanarról a jelenségről teljesen mást gondoljanak a választók bal és jobb oldalon.

A politika nem csak Magyarországon polarizálódik egyre inkább, Trump 2016-os megválasztása óta beléptünk az új tribalizmus korszakába,⁹² ahol a saját tábort kell erősíteni, még akkor is, ha nem értünk vele maradéktalanul egyet, mert a másik tábor sokkal rosszabb, ha ők hatalomra kerülnek (hatalmon maradnak) az felér egy ‘világvé-

⁸⁵ Ináncsi i. m. 16.

⁸⁶ Bajomi-Lázár i. m. 8.

⁸⁷ Çakır-Somlyai i. m. 114.

⁸⁸ Bayer (2022) i. m. 1.

⁸⁹ Barts i. m. 36.

⁹⁰ Wardle–Derekshhan i. m.

⁹¹ ‘Post-truth’ declared word of the year by Oxford Dictionaries. *BBC*, 2016. november 16. <https://www.bbc.com/news/uk-37995600> (saját fordítás P.R.)

⁹² Thomas J. Froehlich: A disinformation-misinformation ecology: The case of Trump. In: Ján Visňovský – Jana Radosinská (szerk.): *Fake News is Bad News-Hoaxes, Half-Truths and the Nature of Today's Journalism*. Intechopen, 2021.

gével'. Emiatt a politikai vezetők egy részén képződik egy 'teflon bevonat' – mindegy mit mondanak, mindegy mit csinálnak, a támogatók szemében a kritika csupán a másik tábor ármánykodása lesz. A demokráciában bízók számára ugyanakkor jó hír lehet, hogy szoros eredmények esetén elég a két táboron kívüli bizonytalan szavazókat sikeresen megszólítani, akik feltehetően sokkal inkább hallgatnak az észérvekre. Például a látszólag áttörő győzelem ellenére a Biden–Trump vagy a Trump–Harris választás is csupán néhány csatátér államban és pár százezer választón dőlt el.⁹³

A *post-truth*, a politikai polarizáció és a tribalizmus mellett látszólag kevésbé meghatározó a *deepfake* elterjedésének egy nem szándékolt következménye, a bizalom csökkenése, aminek valójában sokkal átfogóbb hatása lehet. Vaccari és Chadwick⁹⁴ Egyesült Királyságban végzett kvantitatív kutatása alapján a *deepfake* technológia és a dezinformáció komoly kihívást jelent a hírek hitelessége és a közvélemény bizalma szempontjából is. A mesterséges intelligencia által létrehozott szintetikus videók még akkor is növelhetik a bizonytalanságot és csökkenthetik a hírekbe vetett bizalmat a közösségi médiában, ha nem tévesztik meg a felhasználókat. Mindez hosszú távon káros hatással lehet a demokratikus társadalmakra, mivel megváltoztatja az online kultúrát, kevésbé együttműködő és felelősségteljes viselkedési formákat, illetve normákat hoz létre.

Ha szűken nézzük a generatív MI-t, akkor elmondható, hogy az MI manipulációnak pusztán az elvi lehetősége bármit képes hiteltelessé tenni, mert könnyen rámondják bármire, hogy *deepfake*. Ehhez nem is kell *deepfake*-nek lennie, elég pusztán az a tudat, hogy az lehet. Ugyanígy „[a]z álhírek kifejezés gyakori használata a vezetők – tekintélyelvű és demokratikus vezetők egyaránt – által aláásta a média hitelességét, keretet teremtve a vezetők számára a velük szemben kritikus tudósítások lejáratására.”⁹⁵ Tehát egyelőre nem is annyira magától a *deepfake*-től kell tartani, hanem az általa közvetlenül okozott bizalmatlansági és hitelességi válság miatt.

3. Befejezés

3.1. Korszakváltás?

A dezinformációs lánc egyes lépéseinek vizsgálata után összegezzük az eredményeket, majd térjünk vissza a cikk címében is megfogalmazott dilemmához, azaz, hogy új korszakot hoz-e a mesterséges intelligencia a dezinformációban. Először röviden tekintsük át a dezinformációs lánc lépéseit és az azokkal kapcsolatos főbb jelenségeket, fogalmakat, amelyeket fentebb részletesen is leírtam:

⁹³ Ernesto F. L. Amaral: Popular Vote and Voter Turnout in the 2016, 2020, and 2024 US Presidential Elections. *Center for Open Science*, 2024. <https://doi.org/10.31219/osf.io/gbsdy>

⁹⁴ Cristian Vaccari – Andrew Chadwick: Deepfakes and disinformation: Exploring the impact of synthetic political video on deception, uncertainty, and trust in news. *Social media+ society*, Vol. 6., No. 1. (2020) <https://doi.org/10.1177/2056305120903408>

⁹⁵ Çakır-Somlyai i. m. 114.

1. táblázat: A dezinformációs lánc lépései, kapcsolódó fogalmai és jelenségei

A dezinformációs lánc lépése	Kapcsolódó fogalmak és jelenségek
motiváció	sokrétű motivációk, tényrelativizmus, információs zaj, <i>mockumentary</i> és áldokumentumfilm, manipuláció, tudományos hamisítás
létrehozás	generatív MI, <i>deepfake</i> , Wardle-féle hét lépcsős skála, arc kicserélése, <i>lip syncing</i> , eszközök elérhetőségének demokratizálódása
terjesztés	<i>misinformation</i> (félretájékoztatás), közösségimédia-algoritmusok, moderáció, <i>astroturfing</i> , összehangolt nem hiteles viselkedés, részvételi propaganda, álhírek gyorsabb terjedése, szűrőbuborék, visszhangkamra, álhírgyárok, <i>micro-targeting</i> , információs hadviselés
(detektálás)	gyanús jelek, tényellenőrzés, <i>community notes</i> , folkszonómia, RICHDATA piramis, MI megoldások (terjesztők azonosítása, terjedési mintázat detektálása, tartalom felismerése), <i>shadow banning</i>
befogadás	dezinformáció igazságtartalma, <i>deepfake</i> és dezinformáció ellenőrző kapacitás, Haidt moráleméleti kategorizációja, a befogadás kontextusa, összeesküvés elméletek, közösségi média üzleti modellje, ingyenesség, figyelem fenntartása, érzelem kiváltása, közösségimédia-algoritmusok, ember és gép egyesülése
hatás és következmény	álhírek terjedése, amerikai elnökválasztás, Brexit, demokratikus intézmények aláásása, demokratikus működés megzavarása, hírekbe vetett bizalom csökkenése, online kultúra és viselkedési normák megváltozása, hitelesség megítélése, illuzórikus igazsághatás, <i>post-truth</i> korszak, politikai polarizáció, tribalizmus, média hitelességének aláásása

Mindezek alapján az egyes lépések mentén az MI (korszakalkotó) szerepét illetően az alábbiak látszanak:

- motiváció: a motivációk ugyan nem változnak, de azzal, hogy demokratizálódik a dezinformációra, *deepfake* készítésére alkalmas eszközök elérése, elvileg egyre többen készíthetnek ilyen tartalmakat;
- létrehozás: a hihető dezinformáció létrehozása sokkal egyszerűbbé válik az MI alkalmazásával, de a generatív MI megjelenése óta a várakozások ellenére egyelőre elmaradt a dezinformációs tartalmak robbanásszerű megnövekedése;
- terjesztés: színlelt tömeg (botok és fizetett trollok) segítségével széles körben terjeszthető a dezinformáció, a *microtargeting* segítségével célzottan juttathatók el a dezinformációt tartalmazó üzenetek a befogadókhoz, akik meggyőződésből vagy tévedésből tovább terjeszthetik azokat (részvételi propaganda);
- detektálás: az MI eszközök segítségével több fajta módon is lehet a dezinformáció ellen küzdeni (terjesztők azonosítása, terjedési mintázat detektálása, tartalom felismerése), de ezek egyelőre nem elég megbízhatók és gyorsak;
- befogadás: mivel a befogadás a közönség oldalán történik meg, szűk értelemben véve ebben játszanak egyelőre a legkevésbé szerepet az MI eszközök (azok inkább a létrehozásban, terjesztésben és detektálásban meghatározóak);
- hatás és következmény: ahogy a befogadás esetén, a hatásban sincs közvetlen szerepe az MI-nek, közvetve azonban óriási változásokat indukálnak, ami akár új korszakot is jelenthet a dezinformációban (erről bővebben lásd lentebb).

A fentebbiek alapján tehát a létrehozás, a terjesztés és részben a detektálás terén van egyelőre közvetlen szerepe az MI-nek a dezinformációs láncban. De mindezek révén közvetve mégis a befogadás és a széleskörű társadalmi következmények terén hoz érdemi változást a mesterséges intelligencia megjelenése a dezinformációban. Ahogy Bayer Judit találóan összefoglalja:

„A mesterséges intelligencia szerepe különösen jelentős a dezinformációval való összefüggésben. E technológiai megoldások létezése előtt a dezinformáció terjesztése költséges, munkaigényes és lassú volt. Az algoritmusok, különösen a gépi tanulás alkalmazása felerősítheti a dezinformáció létrehozását, megjelenését és hatókörét, ezáltal hatását is.”⁹⁶

Állításom szerint a mesterséges intelligencia kapcsán a terjesztés és a társadalmi hatás terén tapasztalható érdemi (negatív) változás az MI dezinformációs láncban betöltött szerepében, nem pedig a dezinformáció létrehozásában, ami a generatív MI megjelenése óta az érdeklődés középpontjába került. Hiába lehet ugyanis sokkal hitehetőbb dezinformációs tartalmakat generálni, ha a létrehozók nem lennének képesek érdemben terjeszteni azokat, akkor sokkal kisebb hatást lennének képesek kifejteni. Ráadásul, mint láttuk, az emberek nem azért hisznek el valamit, mert annyira hihető, hanem mert el akarják hinni, s ebben a terjesztés-(detektálás)-befogadás lépéseknek sokkal fontosabb szerepe van, mint magának a létrehozásnak. Ha MI fémjelezte korszakváltásról beszélhetünk, a dezinformációban az nem a generatív MI-nek, hanem egyértelműen a közösségi média MI-algoritmusainak köszönhető. De ez a változás sem ment volna végbe az emberek aktív közreműködése nélkül. Például a detektálásban is kulcsfontosságú a *human-in-the-loop* megközelítés, nem automatizálható teljesen a dezinformáció észlelése és kezelése, moderátoroknak kell felügyelni (verifikálni) az MI eszközök munkáját.⁹⁷ Valójában egy olyan *hibrid dezinformációs láncról* van szó, amelyben maguk az emberek játszanak kulcsszerepet a terjesztésben, detektálásban, befogadásban és a társadalmi változások indukálásában is. A közönség, az influenszerek, celebek és politikusok szinte elválaszthatatlanok a botoktól, fizetett trolloktól és a közösségimédia-algoritmusok működésétől.

Ha korszakhatárt kellene meghatározni a cikkben taglalt jelenségeknek, akkor az 2016 lehetne, amikor a Brexit sokkolta a világot, elnökké választották Trumpot és a *post-truth* kifejezés az év szava lett. Az, hogy ez tartós változás, jól jelzik a COVID-járvány idején terjedő álhírek és az oltásellenesség, illetve a Capitolium 2021 januári ostroma, majd Trump újraválasztása 2024-ben.

De vajon hozhat-e újabb korszakot az MI további fejlődése a dezinformációban? Krekó szerint a generatív MI megjelenéséig alapvetően a terjesztésben játszott fontos

⁹⁶ Bayer (2022) i. m. 1.

⁹⁷ Gianluca Demartini – Stefano Mizzaro – Damiano Spina: Human-in-the-loop Artificial Intelligence for Fighting Online Misinformation: Challenges and Opportunities. *Bulletin of the IEEE Computer Society Technical Committee on Data Engineering*, Vol. 43., No. 3. (2020) 65–74. <https://tinyurl.com/24tz3s4>

szerepet az MI a közösségimédia-algoritmusok révén.⁹⁸ Idővel azonban a tartalomgyártás is automatizálhatóvá válhat és akkor már nem is kell feltétlenül ember hozzá. De erre az esetre is igaz, hogy nem a dezinformáció az érdekes, hanem az általa kifejtett társadalmi hatás, tehát az embereket nem lehet kihagyni a dezinformációból.

Az várható, hogy a jövőben az MI-t még intenzívebben használják majd a dezinformációnak szánt tartalom létrehozásában is. Az MI fejlesztése 2024 óta az önállóbb ágensek felé mutat, amelyek el tudják végezni a rájuk bízott feladatokat egyfajta MI asszisztensként,⁹⁹ így bevethetők akár dezinformáció tömeges kreálására is. Ez megerősítheti a 2016-ra datálható, közösségi médiához köthető korszakváltást, de feltételezően nem hoz létre új korszakot mindaddig, amíg a közösségi médiára fókuszáló terjesztés nem változik érdemben. Az viszont már a tudományon túlmutató, sci-fi felé hajló kérdés, hogy vajon az MI képes lesz-e valaha önállóan, saját szándék alapján tömegesen manipulálni az embereket dezinformációval – ez lehetne ugyanis a következő korszak, de reméljük, hogy ilyesmire soha nem kerül majd sor.

3.2. Mi lehet az MI szerepe a dezinformáció elleni küzdelemben?

Az általam javasolt dezinformációs lánc alkalmazása nem csak abban segíthet minket, hogy könnyebben legyünk képesek azonosítani és megérteni a mesterséges intelligencia szerepét a dezinformációban, hanem hogy azonosíthassuk azokat a specifikus beavatkozási pontokat és lehetséges ellenlépéseket, amivel a dezinformáció ellen *szisztematikusan* fel lehet lépni. Ezek alapján véleményem szerint több szintű a probléma kezelésének a lehetősége:

1. dezinformáció készítéséhez kapcsolódó motiváció gyengítése (pl. a *deepfake porn* készítésének kriminalizálása vagy a kattintás alapú reklámbevételekből élő álhírgyárak monetizációs képességének korlátozása);
2. dezinformáció létrehozásának megnehezítése (pl. generatív MI-platformok felhasználási körének korlátozása), az MI által létrehozott tartalmak vízjelzése – bár sajnos ezeket egyelőre könnyű hamisítani,¹⁰⁰
3. a dezinformáció detektálásának és címkézésének megerősítése, elterjesztése és tudatos használatának elősegítése, a közösségimédia-platformok dezinformáció detektálási gyakorlatának szabályozása;¹⁰¹
4. a dezinformáció terjedésének megakadályozása vagy legalább lassítása, hatókörének korlátozása (pl. moderálás, hátrébb sorolás, *shadow banning* stb.);
5. a dezinformáció befogadásának alakítása, a társadalom felkészítése a dezinformációra egyéni és társadalmi szinten: pl. közös biztonsági szó az unokázó csalók ellen, a tudatos médiahasználat elősegítése, a médiaműveltség oktatása, információs írástudás szintjének javítása;

⁹⁸ Krekó i. m. 96.

⁹⁹ Az OpenAI MI első ágensét 2025 elején jelentette be: <https://openai.com/index/introducing-operator/>

¹⁰⁰ Bayer (2024) i. m. 90.

¹⁰¹ Azgin–Kiralp i. m.

6. ha mégis elterjedt, akkor a dezinformáció negatív hatásának csökkentése.¹⁰²

Mindebben segítségünkre lehet az MI is, többek között a dezinformáció felderítésében és a terjedés korlátozásában.¹⁰³ De a dezinformáció elleni küzdelem feladata elsősorban nem az MI-é, hanem közös ügy: „[a] dezinformáció elleni küzdelem nemcsak a kormányok és a médiumok felelőssége, hanem az egyéneké, a családoké és a közösségeké is.”¹⁰⁴ A felhasználókon túl össze kell(ene) fognia a közösségimédia-platformok tulajdonosainak és fejlesztőinek, a szabályozásért felelős döntéshozóknak, kutatóknak és újságíróknak, hogy megfelelő kereteket dolgozzanak ki a dezinformáció kezelésére a létrehozás, terjesztés és detektálás terén.

Bármilyen lépéseket teszünk is azonban, meg kell tanulnunk együttélni azzal, hogy az MI felerősíti a *deepfake* generálását és terjesztését. De fontos, hogy ettől még ne fatalista állítás legyen az, hogy ‘nem hihetünk a szemünknek és fülünknek’, tehát nem minden eleve hazugság és nem is vagyunk teljesen tehetetlenek ez ellen. Ez ugyanis a *post-truth* korba való belépés kritikátlan elfogadása lenne, ahol bárkinek lehet saját ténye, mivel azt is olyannak tekintik, mint a véleményt és annak van igazsága, aki a leghangosabban és a legkitartóbban kiabál. Ha nem hihetünk a *deepfake* miatt a saját érzékszerveinknek, az még inkább megnöveli a szkepszis szintjét, a mindenben való szisztematikus kételkedés viszont megágyaz az összeesküvés elméleteknek:

„ez egyébként a post-truth korszak egyik nagy paradoxona: a manipuláció ellen éppen az egészséges szkepszis és gyanakvás vértelmezheti fel az újságírókat és az állampolgárokat. Azonban a mindent eluruló, cinizmusba forduló szkepszis ad lehetőséget a közvélemény összeesküvés-elméleteken keresztüli manipulációjára”¹⁰⁵

Tehát a ‘nem hihetünk az érzékszerveinknek’ állítás akkor jótékony hatású, ha mindenben csak egy *egészséges* kételkedést alakít ki. Vagyis ne higgyünk el bármit, csupán azért, mert hihetőnek látszik, ellenőrizzük le az információ kontextusát is. De egyúttal tudjuk azt is, hogy bizalom nélkül nem működik egyetlen emberi társadalom sem.

¹⁰² Krekó Péter a dezinformáció és az összeesküvés elméletek elleni küzdelem négy legfontosabb eszközét az alábbiakban látja: megelőző csapás (előzetes kínálatcsökkentés), válaszcspapás (utólagos kínálatcsökkentés), gyógyítás (utólagos keresletcsökkentés) és immunizálás (előzetes keresletcsökkentés). Krekó i. m. 275. Ezek összhangban vannak az általam a 3–6. pontokban jelzett lehetséges beavatkozásokkal.

¹⁰³ Bayer (2024) i. m. 91.

¹⁰⁴ Çakır-Somlyai i. m. 106.

¹⁰⁵ Krekó i. m. 268.

ÚTMUTATÓ EGY ISKOLAI (CYBER)BULLYING SZABÁLYZAT MEGALKOTÁSÁHOZ

Pongó Tamás*

1. A téma aktualitása és újabb szabályozása

Magyarországon a *bullying* és *cyberbullying* elleni küzdelem 2016 óta a Digitális Gyermekvédelmi Stratégia részét képezi,¹ azonban kifejezett anti-bullying törvény a mai napig nem született. Ez azonban nem jelenti azt, hogy ne történtek volna jogalkotói lépések, még ha azok súlya igencsak csekély.

Az említésre érdemes friss változások közül az első a 2024. évi XXX. törvény, amely több jogszabályt is módosított a cyberbullying elleni védekezés jegyében. Egyrészt, az 1997. évi XXXI. törvényt a gyermekek védelméről és a gyámügyi igazgatásról (Gyvt.), amelynek 6. § (7) bekezdése értelmében ezentúl „a gyermeknek joga van ahhoz, hogy tudatos és felelősségteljes internet-, valamint médiahasználatot – ideértve a zaklatás internetes formáival szembeni hatékony fellépést – biztosító nevelésben, oktatásban és képzésben részesüljön.”² A törvényhez fűzött általános indokolás alapján ezen új szabályok bevezetésének célja, hogy „[a] jövőben nem lehet olyan iskolás gyermek az országban, aki nem tudja hová forduljon segítségért abban az esetben, ha internetes zaklatás áldozatává vált. A javaslat nyomán egyszerűbbé válik a kiskorúak személyiségi jogait sértő, a cyberbullying kategóriáját is kimerítő tartalmak jelentése és azok eltávolítása.”³ Ennél pontosabban azonban a törvény részletes indokolása sem fogalmaz, így nincs könnyű helyzetben a jogalkalmazó, amikor azt kívánja feltárni,

* PhD, egyetemi adjunktus, Szegedi Tudományegyetem, Állam- és Jogtudományi Kar, Közjogi Intézet.
ORCID: <https://orcid.org/0009-0001-1727-5882>

¹ 1488/2016. (IX.2.) Kormányhatározat a Gyermekek Számára Biztonságos Internetszolgáltatás megteremtéséről, a tudatos és értékteremtő internethasználatról és Magyarország Digitális Gyermekvédelmi Stratégiájáról.

² 1997. évi XXXI. törvényt a gyermekek védelméről és a gyámügyi igazgatásról. A módosítást a 2024. évi XXX. törvény a gyermekek védelme érdekében egyes törvények módosításáról 5. §-a vezette be.

³ 2024. évi XXX. törvény indokolás a gyermekek védelme érdekében egyes törvények módosításáról szóló 2024. évi XXX. törvényhez.



hogyan a mindennapi életben mit jelent ez a fajta hatékony fellépés és felelősségteljes használat.⁴ Ráadásul felmerül a kérdés, hogy vajon mire alapozza az indokolás azt az állítást, hogy a Gyvt.-be bevezetett rendelkezés hatására egyszerűbbé válik majd egy adott platform jelentéstételi és tartalomeltávolítási rendszere. A törvény módosítása értelmében a gyermek képzést kap, amely tudás birtokában könnyebben felismeri, hogy egy adott közösségimédia-tartalom sérti a személyiségi jogait, és pontosan megismeri a jelentéstételi folyamatot is, így szükség esetén sikeresen jelenteni tudja a sértő tartalmat az adott platformon. Eddig valóban segítséget kíván adni az új módosítás. Ám az továbbra is rejtély, hogy ettől a jelentés technikai folyamata, és főként maga az eltávolítás miként válna egyszerűbbé. Jóindulatúan szemlélve arra lehet csupán gondolni, hogy e sorokat a módosító törvényben később megfogalmazottakkal összhangban kell értelmezni. Ugyanis a 2024. XXX. törvény módosította az elektronikus hírközlésről szóló törvényt is, amelynek nyomán a Nemzeti Média- és Hírközlési Hatóság (továbbiakban NMHH) feladata a kiskorúak és szüleik segítése az „internetes bántalmakkal, visszaélésekkel és jogsértésekkel szembeni fellépésben”⁵ tudatosító programok és kampányok útján. Ennek része az NMHH által működtetett Internet Hotline segítségnyújtó szolgáltatás is, amely kifejezetten a biztonságos online környezet kialakításáért felel.⁶ Következésképpen, ha a gyermek tudatos internethasználóvá válik, akkor ismeri az Internet Hotline-t, tudatában van annak, hogy a Hotline kiemelt jelentéstételi szerepkörrel (*trusted flagger*) bír, így az itt jelentett tartalom vélhetően egyszerűbben és gyorsabban eltávolítható.

További fontos változás, hogy a Gyvt. módosításával a gyermekeknek kötelezettsége a véleménynyilvánítás szabadságának gyakorlása során a többi gyermek emberi méltóságának tiszteletben tartása, beleértve az internet- vagy médiaszolgáltatás útján történő véleménynyilvánítást is.⁷ A módosítások szövegét, valamint az általános és részletes indokolásban foglaltakat olvasva látható, hogy a jogalkotás egy emberi méltóságon alapuló, személyiségi jogi kérdésként való értelmezés irányába tolja el a jogi értékelést. Ugyanis a kiskorú személyiségi jogait sértő tartalom eltávolítása a cél, miközben a gyermekeknek kifejezett kötelezettségévé tették a másik gyermek emberi méltóságának tiszteletben tartását az online, mediatizált közegben.

Az online térben való emberi méltóságon alapuló jogvédelem vonalát erősítő jogalkotói szándékot jelez a második friss jogszabályi változás, nevezetesen a T/9996. számú törvényjavaslat 3. §-a, amely egy új büntetőjogi tényállását vezetett be, az internetes agressziót. E szubszidiárius bűncselekményi tényállás a következő:

„(1) Aki nagy nyilvánosság előtt elektronikus hírközlő hálózat útján olyan kifejezést, ábrázolást vagy kép- és hangfelvételt használ vagy tesz közzé, amely beazonosítható személlyel vagy személyekkel szembeni erőszakos a) halált

⁴ 2024. évi XXX. törvény indokolás 5. §. 7. §.

⁵ 2024. évi XXX. törvény indokolás 44. §.

⁶ Balla Boróka – Bartók Zita – Csalár Dorina: A gyermekvédelem és tartalomszabályozás új útjai az interneten – Az Internet Hotline működése és gyakorlata. *In Medias Res*, 2023/1. 138–168.

⁷ A Gyvt. 10. § (1) bek. d) pontot módosították a 2024. évi XXX. törvény indokolás 5. és 7. §-ai.

okozó, vagy b) különös kegyetlenséggel elkövetett büntetendő cselekményre irányuló szándékot vagy kívánságot fejez ki...⁸

A tényállás szövege és az ahhoz fűzött indokolás elemzése alapján kijelenthető, hogy a gyermekek által elkövetett cyberbullying magatartások egy része megvalósítja ezeket a tényállási elemeket, így azok büntetendők. Példaként említeném az első széles körben ismertté vált cyberbullying halálesetet, amely Megan Meier nevéhez kötődik. Megan egy nyilvános MySpace kép alá kapott durva hangvételű kommenteket, mint például „jobb lenne a világ, ha meghalnál”, és ez is hozzájárult ahhoz, hogy Megan öngyilkosságot követett el.⁹ Mindez 2006-ban, az Egyesült Államokban történt, ami bizonyítja számunkra, hogy milyen régóta él velünk ez a probléma. Hazánkban az első hasonlóan nagy médiavisszhangot kiváltó tragédia Tisza Gábor nevéhez köthető, aki szintén bullying miatt követett el öngyilkosságot 2021-ben.

Az új hazai tényállás értelmében, ha egy 16 éves gyermek (büntetőjogi értelemben fiatalkorú) valamely közösségi média platformon megosztott képe, gif-je vagy kommentje valakinek a halálát vagy egyéb, különösen kegyetlen cselekmény áldozatává válását kívánja, úgy a fiatal bűncselekményt követ el. Az indokolás szerint „az érintett személyek és a köz sérelmének a valós veszélyétől és a sértett szubjektív felfogásától függetlenül is megvalósul a bűncselekmény.”¹⁰ Jelen tanulmánynak nem célja elemezni, hogy e büntetőjogi tényállás megfelelően védi-e a jogi tárgyat, és megfelel-e a büntetőjogi dogmatika minden elemének (bár ezzel kapcsolatban aggályok merülnek fel). Csupán illusztrálni kívánta, hogy az online térben megvalósuló, emberi méltóságot sértő magatartások jogalkotás útján való rendezése egy akut és aktuális problémakör, amelynek egy aspektusához kíván kiindulási alapot nyújtani ezen írás, egy közoktatásban alkalmazható bántalmazás elleni minta-szabályzat kimunkálásával.

2. Az offline és online bántalmazás jelenléte a hazai diákság körében

Sajnálatos módon, a bullying és cyberbullying jelensége Magyarországon is veszélyezteti a gyermekek mindennapjait. Hazánkban több kutatás is igyekezett feltárni ennek mértékét, noha az adatok vizsgálata után a probléma mértéke pontosan nem meghatározható, de annak jelenléte az iskolai közegben bizonyítható.

Az Unicef a Telenorral közösen, 2014 őszén egy nem reprezentatív kutatást végzett 10-19 éves korú diákok körében (N=1191 fő), amely alapján megállapították, hogy minden harmadik gyermeket ért már pszichológiai az online térben, de csupán minden tizedik kért segítséget.¹¹ 2022 őszén szintén az Unicef megbízásából a Medián Közvélemény- és Piackutató Intézet készített felmérést a magyar középiskolások között

⁸ T/9996. számú törvényjavaslat 3. §.

⁹ Raphael Cohen–Almagor: Cyberbullying, Moral Responsibility, and Social Networking: Lessons from the Megan Meier Tragedy. *European Journal of Analytic Philosophy*, Vol. 16., No. 1. (2020) 83–84.

¹⁰ T/9996. számú törvényjavaslat 3. §-hoz fűzött részletes indokolás.

¹¹ Hazai kutatások – A gyerekek közel fele szerint egy pofon befelé. Unicef, 2014. <https://tinyurl.com/3yx9ubeu>

(N=1000 fő). Az eredmények tanúsága szerint, a kitöltők 85%-t érte offline iskolai bántalmazás, s ezen bántalmazotti kör felét (49%) az online világban is érte atrocitás.¹² A cyberbullying vonatkozásában a kitöltő gyerekek 44%-a állította, hogy online bántalmazás áldozata volt a tanulmányai alatt. Ennek legjellemzőbb fórumai a privát üzenetek (67%), valamint az Instagram (34%) és Facebook (26%) platformok voltak.¹³

2017-ben az NMHH a Psyma Hungary Kft.-vel közösen indította el a Digitális Szülőség nagymintás empirikus kutatást. Annak keretében 7-16 éves (N=2000 fő) gyermekeket és szüleiket kérdezték a médiahasználatról.¹⁴ A felmérést eddig kétszer végezték el, először 2017-ben, majd 2020-ban, s ennek keretében a cyberbullying előfordulását is vizsgálták. Az eredmények alapján a 7-16 éves korosztály 7%-a tapasztalt online bántalmazást 2017-ben, és 9%-uk 2020-ban. Noha ez biztató eredménynek tűnhet, mégis jellemző, hogy – a 9-10 éves korcsoportot kivéve – minden korosztálynál növekvő tendenciát mutat az áldozattá válás.¹⁵

Az NMHH egy másik, cyberbullyingra irányuló kutatása a 10-18 éves magyar diákok érintettségét vizsgálta. A kutatás alapjául egy országos reprezentatív kutatás magyar diákokra vonatkozó adatai szolgáltak (*Health Behaviour in School-aged Children*, azaz Iskoláskorú Gyermekek Egészségmagatartása). Azokból másodelemzéssel és fókuszcsoportokon keresztül készítettek elemzést.¹⁶ A nagyjából 6000 fős minta alapján az áldozatok aránya 17,8% volt, míg az elkövetőké 12,5%.¹⁷ Az összes válaszadó 23,3%-a érintett valamilyen cyberbullying magatartásban, 7%-uk pedig mind az áldozati, mind az elkövetői oldalon részt vett bántalmazásban.¹⁸

A HBSC 2022-re vonatkozó nemzeti jelentése szerint, (N = 6242 fő) az offline bullying a gyermekek kb. negyedét érintő probléma.¹⁹ A diákok 10,6%-a rendszeres bántalmazások áldozata, ám az életkor előrehaladtával az áldozattá válási gyakoriság egyre csökken.²⁰ A cyberbullying áldozattá válási adatok is biztatóak, hiszen egy-két alkalommal a diákoknak csupán 13,8%-a, míg rendszeresen 5,3%-uk esett áldozatul.²¹

A fent említett statisztikai adatok mellett fontos kiemelni, hogy több, validált mérési eszköz is elérhető, így a hazai bullying és cyberbullying magatartások tudományos alapú vizsgálata biztosított.²²

¹² Középiskolások tapasztalatai az iskolai zaklatásról. Unicef–Medián, 2022. <https://tinyurl.com/4pvr4wew>

¹³ Uo.

¹⁴ Digital Parenting Kutatás 7-16 éves gyermekekkel és szüleikkel – Kutatási jelentés a Nemzeti Média-és Hírközlési Hatóság részére. 2–3. <https://tinyurl.com/3kp732dd>

¹⁵ Digital Parenting Kutatás i. m. 147.

¹⁶ Jármí Éva – Várnai Dóra Eszter: Online platformok és a bullying megjelenési formái. *Online platformok*, 2023. július 13. 4. <https://tinyurl.com/5ebenpjr>

¹⁷ Uo. 7.

¹⁸ Uo. 7.

¹⁹ Németh Ágnes (szerk.): *Iskoláskorú gyermekek egészségmagatartása 2022*. Budapest, ELTE, Pedagógiai és Pszichológiai Kar, 2023. 57. <https://tinyurl.com/atm2ytw>

²⁰ Uo. 58.

²¹ Uo. 60.

²² Validált mérési eszközökhöz ld. Nikolett Arató: *The Role of Socio-emotional Skills in Cyberbullying Engagement*. Doktori értekezés. Pécs, Pécsi Tudományegyetem, Pszichológia Doktori Iskola, 2021. 2.

Áttekintve a fentebbi adatokat, látható, hogy az áldozattá válási mutatók rendkívüli eltérést mutatnak. Különösen szembeötlő az Unicef által mért adatok összevetése a HBSC eredményekkel, illetve az arra épülő NMHH kutatással. Míg az Unicef adatai közel 30-45% között áldozati kört vázolnak fel, addig a nagymintás nemzeti jelentés csupán ennek töredékét, nagyjából 10% körüli értéket állít. E nagyfokú eltérések még inkább indokolják további empirikus kutatások elvégzését.

Összefoglalóan: noha a statisztikai adatok eltérést mutatnak a probléma intenzitását, valamint az áldozati kör mértékét illetően, azt egyértelműen bizonyítják, hogy a bántalmazás offline és online formája jelen van a hazai oktatási környezetben. Ezért sürgető, hogy készüljön Magyarországon is államilag támogatott anti-bullying program, hiszen a Tabby in Internet vagy a KiVa program is pénzügyi okokból szűnt meg. Jelen tanulmány ebben kíván segítséget nyújtani az oktatási intézményeknek egy mintaszabályzat kidolgozásával. Ez természetesen csak kiindulási alapként szolgál, és támogatást nyújt azon iskolák számára, akik tenni kívánnak egy biztonságosabb oktatási környezetért.²³

3. A köznevelésben bevezethető megfélemlítés elleni szabályzat ajánlott elemei

A tanulmány jelen része egy köznevelésben alkalmazható bántalmazás elleni mintaszabályzat főbb elemeit mutatja be, amelyek a következők: hatály; definíciók; vétségek; jelentéstétel; eljárás és szankciók.

3.1. Hatály

Egy szabályzat megfelelő alkalmazásához pontosan meg kell határozni annak hatályát (alkalmazási körét). Tekintettel arra, hogy a fókuszban a köznevelés áll, így a Köznev. tv. által kijelölt intézményekre²⁴ és ahhoz tartozó személyekre kell kiterjednie a szabályzat hatályának. Az intézményi keretek között jelen tanulmány az óvodai oktatást és a felnőttképzést nem tárgyalja. A személyi hatály kérdését fontos tisztázni, hiszen a Köznev. tv. alapvetően hárompólusú jogviszonyként határozza meg a nevelés kérdését: középpontjában a tanuló, a pedagógus és a szülő állnak. Természetesen, a tanuló és a pedagógus szabályzat alá tartozása egyértelmű, bár fontos pontosítani, hogy a köznap értelemben vett tanáron kívül minden olyan személy a szabályzat hatálya alá tartozik, aki valamilyen módon érintkezik a gyermekkel. A törvény 4. §-ában definiált nevelőtestület minden tagja, az óraadó, az utazó gyógypedagógus, az utazó konduktor, az együttnevelést segítő pedagógus, pedagógiai asszisztens mellett a nem-tanító személyzet (pl. iskolatitkár, iskolapszichológus, karbantartó, a portás vagy éppen a konyhai al-

fejezet, 2.2. és 2.3. alfejezetek.; Margitics Ferenc – Figula Erika – Pauwlik Zsuzsa – Szatmári Ágnes: A szülői bánásmód hatása az iskolai erőszakkal kapcsolatban előforduló magatartásmintákra. *Magyar Pedagógia*, 2010/3. 211–238.

²³ Gyakorlatban megvalósult bántalmazás elleni fellépésekért ld. Kulcsár Gabriella: *Bullying és cyberbullying – Hogyan előzzük meg és kezeljük az iskolai és az online kortársbántalmazást?* Budapest, Novissima, 2024.

²⁴ 2011. évi törvény a nemzeti köznevelésről 7. § (1) bek.

kalmazott) is. E személyi kör könnyen azonosítható, hiszen mindenki beletartozik, aki bármilyen jogviszony keretében érintkezik vagy érintkezhet a gyermekkel, amíg ő az intézmény nevelése-felügyelete alatt áll, tehát a 20/2012. (VIII. 31.) EMMI rendelet 1. § c) megfogalmazásában együttesen az alkalmazottak. Ugyanakkor a személyi hatály szülőre való kiterjesztését is szükséges megvizsgálni, hiszen a szülő nem áll közvetlen jogviszonyban az oktatási intézménnyel, hanem az ott tanuló (nevelt, örökbefogadott, gyámolt) gyermeke törvényes képviselőjeként vesz részt az iskola életében. Mégis, egy szülő magatartása is megvalósíthat bullying vagy cyberbullying eseményt, például, ha az áldozat egyik szülője bosszúból bántalmazza a gyermeket bántó osztálytársat az online térben. További, a gyermekkel érintkező személy lehet még az önkéntes munkát végző vagy más okból az iskolai tevékenységet segítő civil, egyházi, gazdasági vagy önkormányzati szervezet képviselője, segítője (pl. vetélkedő, kulturális műsor, tájékoztató rendezvény résztvevője, szervezője, adományozó, korrepetáló felnőtt, volt diák), ha az iskola területén tartózkodik, vagy az intézmény működésével kapcsolatba kerül, alkalmilag vagy rendszeresen, nevesített jogviszony nélkül is. Szerencsére az EMMI rendelet megoldást nyújt e kérdésben, hiszen egyrészt az 1. § d) pontjában kiterjeszti a rendelet hatályát a szülőkre, törvényes képviselőkre, másrészt a 4. § (4) bekezdése alapján „[a]z SZMSZ-ben meghatározhatók azok a nevelési-oktatási intézmény biztonságos működését garantáló szabályok, amelyek megtartása kötelező az intézmény területén tartózkodó szülőknek, valamint az intézménnyel kapcsolatban nem álló más személyeknek.” A rendelet 4. § (2) f) pontja pedig lehetőséget ad az oktatási intézményeknek, hogy SZMSZ-ben a „a gyermekek, tanulók egészségét veszélyeztető helyzetek kezelésére irányuló eljárásrendet” fogalmazzon meg. Mindezek fényében, egy anti-bullying szabályzat hatálya, a gyermekek és az alkalmazottak mellett, a szülők magatartására is kiterjeszthető.

A személyi hatályra épül a további alkalmazási kérdések rögzítése. Miről lehet szó? Például, hogy az iskolán kívüli (*off-campus*) magatartásokra kiterjed-e a szabályzat hatálya. Alapvetően az iskolai (*on-campus*) cselekményeknek az intézmény hatáskörébe való tartozása nem vitatott, így az ilyen magatartásokra, eseményekre mindenképpen vonatkozik a szabályzat. Továbbá, az oktatási-nevelési intézményhez kötődő (általa szervezett vagy felügyelt) esemény vagy program (pl. osztálykirándulás) is iskolai programként kezelendő. Ám az iskolán kívüli online magatartásokat nehéz úgy szabályozni, hogy az valóban alkalmazható legyen a gyakorlatban. Ebből kifolyólag e minta-szabályzat a közös minimum-garancia megteremtése érdekében nem terjeszti ki az iskola hatáskörét az intézményhez nem kötődő online magatartásokra. Ugyanakkor ezek jelentős hatás gyakorolhatnak az érintettek iskolai életére, így a máshol már bevált vagy megpróbált megoldásokat²⁵ érdemes egy meghatározott időre, azaz korlátozott időbeli hatállyal bevezetni, majd azt követően értékelni az alkalmazhatóság és visszatartó erő szempontjából. Ezért a hatálynál célszerű megállapítani, hogy a szabályzatot elfogadó entitás mennyi idő után fogja kötelezően felülvizsgálni, hatályában fenntartani vagy megfelelően módosítani az addigi rendelkezéseket.

²⁵ Egy megoldási javaslat olvasható: Pongó Tamás: *A cyberbullying és a diákok véleménynyilvánítási szabadsága, különös tekintettel az USA jogrendszerére*. Szeged, Jurisperitus, 2021. III. fejezet.

3.2. Definíciók

Kiemelt figyelmet kell fordítani a szabályzat megszövegezésére és szóhasználatára. Egyrészt, mindent nagyon pontosan kell megfogalmazni, hiszen önmagában a magatartás definiálása is kulcskérdés, majd azt követően a prevenció és a beavatkozásnál alkalmazott fogalmi halmaznak is zárt rendszert kell alkotnia. Továbbá, figyelemmel kell arra is lenni a megalkotás folyamatában, hogy egy anti-bullying szabályzat nem szükségképpen önálló szabályzatként jelenik meg, hanem az intézmény szervezeti és működési szabályzatába vagy házirendjébe beépítve. Ebben az esetben különösen fontos, hogy a köznevelés jogszabályi környezete által kimunkált fogalmi és logikai rendszerbe organikusán beleilleszthető legyen a bullying-specifikus kifejezésekkel együtt, mint például bántalmazó, szemlélődő. Ezért szerepel például a 3.1 (hatály) pontban az 'alkalmazott' kifejezés. Leírhatóak lennének az iskolai közeggel kapcsolatban álló személyek más módon is, de az zavart okozna, eltérve a köznevelés jogi környezetében már bevezetett fogalmaktól, míg a példában vett 'alkalmazott' kifejezés átfogja azon személyi kört, akikre a szabályzat is rendelkezéseket kíván adni.

Definiálni szükséges, hogy mely magatartás tekintendő iskolainak/iskolán kívülinek, s ezért bizonyos tapasztalatok alapján mérlegelheti az intézmény, hogy a szokásos iskolai közegen kívülre kívánja-e kiterjeszteni a szabályozási (fegyelmezési) jogkörét, azzal a minta-szabályzat 11. §-ban szereplő felhívással, hogy a rendszeres felülvizsgálat módot nyújthat a kiterjesztésre.

Amennyiben e fogalmi elemeket tisztázza a szabályzat, úgy a két legfontosabb kulcsfogalmat szükséges pontosan meghatározni: a bullyingot és a cyberbullyingot.

A bullying már széles körben elfogadott fogalmi elemekkel bír, amelyek az ismétlődés, szándékos károkozás és az erőegyensúly felborítása.²⁶ Természetesen, számos definíció létezik, azonban jelen tanulmány (és az ennek alapján elkészített minta-szabályzat is) a következőt alkalmazza:

'ismétlődő és károkozó magatartás bármely tanuló vagy alkalmazott által, illetve annak sérelmére fizikai és/vagy lelki sérelem, tulajdonában bekövetkező kár okozása vagy arra irányuló szándék, amely erőegyensúly felborulást hoz létre. Fizikai és/vagy lelki sérelem okozása, így különösen a bántalmazott neveltségessé tétele, megalázása, lealacsonyítása vagy dehumanizálása.'

²⁶ A/HRC/31/20, Office of the UN Special Representative of the Secretary-General on Violence against Children, 'Annual report', 5 January 2016. 12. point 61.; Dan Olweus: Invited Expert Discussion Paper – Cyberbullying: An overrated phenomenon?, *European Journal of Developmental Psychology*, Vol. 9. (2012) 523.; Sameer Hinduja – Justin W. Patchin: *Bullying Beyond the Schoolyard – Preventing and Responding to Cyberbullying*. (2nd ed.) Thousand Oaks California, Corwin, 2015. 12.; Nicole L. Weber – William V. Jr. Pelfrey: *Cyberbullying – Causes, Consequences, and Coping Strategies*. LFB Scholarly Publishing LLC, 2014. 9.; Shaheen Shariff: *Sexting and Cyberbullying – Defining the Line for Digitally Empowered Kids*. Cambridge University Press, 2015. 8–9.; Philip C. Rodkin – Karla Fischer: Cyberbullying from Psychological and Legal Perspectives. *Missouri Law Review*, Vol. 77., (2012) 622–626.

Az ismétlődés és az erőegyensúly felborulás az elkövető és az áldozat között szükségserű elemei az offline bullyingnak. Természetesen, a fogalomban használt egyes következmények, úgy mint a fizikai vagy lelki sérülés, vagy tulajdonban történő károkozás, már a törvény és köznevelés által is védett értékek. Mégis, fontos ezek beemelése a bullying definícióba is, hiszen ezáltal illeszkedik egy anti-bullying szabályzat a már létező keretrendszerbe. Továbbá, egy bántalmazás elleni szabályzat célja egy eszköz biztosítása a közösség biztonságának fenntartására, amelynek következtében előfordulhat, hogy egy magatartást adott esetben több szabályzat/eljárás is lefed, így lehetőség nyílik több megoldás közül kiválasztani a legmegfelelőbbet. Például, ha a diák magatartása a fegyelmi eljárás és az anti-bullying szabályzat hatálya alá is tartozik, úgy fennáll a választás lehetősége, vagy akár mindkét szabályzat párhuzamos alkalmazása.

A fentiekben ismertetett tartalmi elemekre épülhet a másik kiemelt definíció, nevezetesen a cyberbullying. E szabályozási kontextusban a cyberbullying

„a diák vagy iskolai alkalmazott által, vagy sérelmére, elektronikus eszköz által elkövetett, szándékos, egyszeri vagy ismétlődő magatartás, amelynek célja az erőegyensúly felborítása, s megvalósítja az alábbiak bármelyikét:

- a) testi-lelki egészség sérelme vagy tulajdonban való károkozás, illetve ezek bekövetkezése a körülmények alapján valószínűsíthető; vagy
- b) ellenséges oktatási környezet létrehozása, vagy az iskola által nyújtott szolgáltatásoktól, hasznoktól és lehetőségektől való megfosztás.”²⁷

Megjegyzendő, hogy a cyberbullyingnak, az offline bullyinggal ellentétben, nincs széles körben elfogadott definíciója.²⁸ Ennek következtében, szükség volt átfogó kutatásra, ami megállapította, hogy melyek a fogalom kulcselemei jogalkotói, jogalkalmazói és akadémiai oldalról. Tekintettel arra, hogy a legszélesebb jogalkotás és joggyakorlat az Egyesült Államokban volt elérhető, így a szerző ezen forrásokra támaszkodva munkálta ki a fent olvasható cyberbullying fogalmat. Első lépésként, minden USA tagállam anti-bullying szabálya került görcső alá, majd a vizsgálat leszűkült azon huszonhét tagállamra, amely *expressis verbis* alkalmazta a cyberbullying terminust. Ennek eredményeként megállapításra kerültek a legfontosabb jogalkotói elemek. Másodikként az Egyesült Államok Szövetségi Legfelső Bírósága (Tinker,²⁹ Fraser,³⁰ Hazelwood,³¹ Morse³²) és más szövetségi és tagállami bírósági döntések (Snyder,³³

²⁷ Pongó i. m. 64.

²⁸ Policy Department for Citizen's Rights and Constitutional Affairs: Cyberbullying Among Young People. European Union, 2016. 9., 19., 26.

²⁹ Tinker v. Des Moines School District, 393 U.S. 503 (1969).

³⁰ Bethel School District v. Fraser, 478 U.S. 675 (1986).

³¹ Hazelwood School District v. Kuhlmeier, 484 U.S. 260 (1988).

³² Morse v. Frederick, 551 U.S. 393 (2007).

³³ JS Ex Rel. Snyder v. Blue Mountain School District, 650 F. 3d 915, Court of Appeals, 3rd Circuit (2011).

Layshock,³⁴ Kowalski,³⁵ Wisniewski,³⁶ J.C.³⁷) elemzése rajzolta ki a jogalkalmazói oldal által kimunkált definíciós elemeket. E két megközelítés összevetésekor felfedezhetővé váltak eltérések a jogalkotói és jogalkalmazói elemek között, így azokat egymásra vetítve formálódott a definíció.

Végezetül, az akadémiai világ fogalmi elemeit tárta fel a kutatás, világszerte. Ellentmondások itt is jelentkeztek, például az erőegyensúly felborulás kötelező eleme minden tudományos cyberbullying fogalomnak, míg az amerikai jogalkotók esetében csupán hatszor volt ezen elem a definíció része,³⁸ sőt Nevada ki is törölte a jogszabályából ezt a jellemzőt.³⁹

A tanulmány által fentebb ismertetett cyberbullying fogalom e három megközelítés (jogalkotói, jogalkalmazói, akadémiai) összevetésének eredményeként született meg. Kiemelendő, hogy minden fogalomnak tartalmaznia kell az erőegyensúly felborulást, mivel a tudományos világ ebben egyöntetű álláspontot képvisel. Továbbá, akár egyetlen cselekmény is megvalósíthat online bántalmazást (pl. egy fénykép feltöltése, ami kommentáradatot indít el), miként azt néhány vizsgált tengerentúli törvény is lefekteti.⁴⁰ A bevezető részben már említett új magyar büntető törvényi tényállás is ezt a nézőpontot támasztja alá. A tanulmány által felvázolt minta-szabályzat szerint az 'elektronikus kommunikáció magában foglalja, de nem kizárólag, a telefonhívásokat, szöveges üzeneteket (SMS), e-maileket, közösségi médiát és az azonnali üzeneteket az online térben.'

Összefoglalóan, a felsorolt definíciós összetevők alkotják a legfontosabb (cyber) bullying definíciós elemeket, amelyeket minden szabályzat fogalom-meghatározási részében szükséges rögzíteni.

3.3. Tiltott magatartások

Egy bántalmazás elleni szabályzat szintén érdemi eleme, hogy milyen magatartást akarunk elítélni, kiszorítani az oktatási intézményekből. A legsúlyosabb vétség, ha egy adott személy a szabályzat hatálya alá tartozó bullying vagy cyberbullying magatartást követ el. Másképp fogalmazva, ha egy diák vagy alkalmazott *követi el ugyanezen személyek bármelyikének sérelmére, iskolában vagy iskolán kívül, de ahhoz kötődő (támogatott, szervezett, felügyelt) eseményen vagy rendezvényen, vagy iskolán kívül* (ha ezt a szabályzat tartalmazza).

³⁴ Layshock v. Hermitage School District, 593 F. 3d 249, Court of Appeals, 3rd Circuit (2010).

³⁵ Kowalski v. Berkeley County Schools, 652 F. 3d 565, Court of Appeals, 4th Circuit (2011).

³⁶ Wisniewski v. Board of Education of Weedsport Central School District, Court of Appeals, 2nd Circuit (2007).

³⁷ J.C. v. Beverly Hills Unified School District, United States District Court, Central District of California, CV 08-03824 SVW (2009).

³⁸ Pongó i. m.

³⁹ Nevada (SB504)(2015).

⁴⁰ California Education Code §48900 (r)(1); Illinois Compiled Statutes 105 ILCS 5/27-23.7 (b).

Továbbá, a tilalmazott magatartások között a megtorlást mint vétséget is indokolt bevezetni, különösen akkor, ha a szabályzat lehetőséget biztosít anonim jelentéstételre (ami javasolt, lásd alább). Kiemelten fontos, hogy minden érintett tisztában legyen azzal, hogy megtorlásnak minősül az a magatartás, amely a bántalmazás elleni szabályzatban foglalt megsértése miatt jelzéssel élő, eljárást indító vagy az eljárásban közreműködő tanulóval, szülővel vagy alkalmazottal szemben ezzel összefüggésben sérelmet okoz, sérelem okozására irányuló magatartást tanúsít, vagy azzal fenyeget. A megtorlás gyakori példája a hamis jelentéstétel, amikor a bántalmazó az ellene tett jelzés retorziójaként hamis jelzéssel él. A megtorlás önálló vétségként való kiemelése elősegíti a hamis vagy rosszhiszemű jelentéstételi gyakorlat megszüntetését az intézményben.

Végezetül, elvi élel szükséges kijelenteni, hogy a szabályzatban foglalt vétségek kizárólagos célja a biztonságos és befogadó oktatási környezet és közösség fenntartása, míg az alkotmányos védelmet élvező véleménynyilvánítások elnyomása nem lehet célja vagy eredménye.

3.4. Jelentéstétel és eljárás

Az oktatási-nevelési intézményben elsődlegesen a megelőzés fontosságát kell hangsúlyozni. Az anti-bullying szabályzatnak elérhetőnek kell lennie az iskola hivatalos weboldalán és minden egyes épületében kinyomtatva is ki kell helyezni. Továbbá, a frissen beiratkozókknak kötelező részt venniük a bántalmazás elleni tréningen az iskolai oktatás első hetében. Az alkalmazottaknak szintén képzésen kell részt venniük, amelyet évente meg kell ismételniük. Ezt részben a tapasztalatok átadása, részben az esetleges szabályok módosulása indokolja. A képzéseknek köszönhetően a közösség minden tagja tisztában lesz a szabályzat létezésével, annak tartalmával, hatályával, a jelentéstételi rendszer jellegével és az eljárással. Mindemellett, az állítólagos bullying vagy cyberbullying magatartás jelentésére több fórumot is biztosítani kell, amit a képzésen is ki kell emelni. Így például online jelentéstételi űrlapot célszerű létrehozni, de telefonon vagy e-mail útján is érdemes fogadni a panaszokat, illetve az iskola által kijelölt személynél vagy szervezeti egységnél is biztosítani kell a lehetőséget. Jelen tanulmány erősen támogatja az anonim módon történő jelentéstételi formát, mivel ez biztonságot nyújt a jelentéstevőknek. Hátránya, hogy bárki bárhonnán jelentést tehet, akár megtorló szándékkal, de ez kezelhető, ha az eljárás hivatalos elindítása előtt, a kijelölt iskolai alkalmazott vagy csapat meghívja egy informális megbeszélésre az érintett diákokat és/vagy alkalmazottakat. Ennek során hatékonyan fel tudják térképezni a helyzetet, és következtetni tudnak arra, hogy miért történhetett a jelentéstétel. Mindazonáltal, az anonim jelentéstétel lehetősége támogatandó, hiszen így a szemlélődők (*bystander*) bátrabban hívják fel az intézmény figyelmét az esetleges aktív bullying vagy cyberbullying magatartásokra anélkül, hogy utána bevonnák őket a formális eljárásba, így megelőzik az esetleges későbbi áldozattá válásukat is.

Természetesen az esetek jelentős részében bizonyítékok összegyűjtésére van szükség. A bullying helyzetek kezelésével megbízott személy/csoport, a bizonyítékok összegyűjtését követően jelentést készít, amelyet kizárólag tájékoztatás jelleggel megküld az intézmény vezetőségének és a fenntartójának. Az ügy érdemben való elbírálására

célszerű az intézmény vezetőségét kijelölni, de megfelelő garanciákat beépítve a döntéshozatali folyamatba, mivel nem zárható ki, hogy a vezetés inkább az esemény eltitkolásában, bagatellizálásában érdekelt. Például a jelentést készítő személy vagy csapat szakmai álláspontját magában foglaló jelentéstől teljesen eltérő, azzal szembenálló döntés nem hozható, illetve összeférhetetlenségi szabályok is kialakíthatóak. Indokolt a döntéshozatalba legalább egy olyan diákot is bevonni, aki semmilyen módon nem érintett a kialakult helyzetben és objektíven képes képviselni a diákság véleményét a helyzet megítélése kapcsán. Bár a diákok bevonásakor mindig érdemes megfontolni, hogy az ügy részleteinek megismertetése egy diáktárssal nem jár-e a sértett érdekeinek sérelmével. Ennek elkerülése érdekében más diákok bevonása csak a sértett egyértelmű hozzájárulása esetén javasolt. Az eljárás nem nyilvános, kizárólag az eljárásban érintett felek és az iskola vezetősége által meghívott személyek vehetnek rajta részt, és köti őket a személyes adatok kezelésének rendje (amelyre az SZMSZ-ek és házirendek általában utalnak). A törvényes képviselőtől túl egyéb képviselőt (pl. ügyvéd megbízása) nem látható el sem a bántalmazó, sem az áldozat oldaláról, hiszen ez nem bírósági eljárás.

A bizonyítékok benyújtását és értékelését követően a döntést szótöbbséggel indokolt meghozni. Ha bármely fél elégedetlen a döntés kimenetelét illetően, úgy jogorvoslattal élhet a köznevelési rendszeren belül. Amennyiben a jogorvoslati eljárás eredményével nem elégedett valamely fél, úgy az általános szabályok szerint lehetősége nyílik peres úton rendezni a helyzetet vagy az oktatási jogok biztosához fordulni.

3.5. Jogkövetkezmények, szankciók

A szankciók elsődleges célja nem a büntetés, hanem az oktatás és a nevelés. Természetesen, egy bántalmazás elleni szabályzatnak tartalmaznia kell komolyabb jogkövetkezményeket is, hiszen a súlyosabb magatartásokra megfelelő választ kell adnia. Ugyanakkor az eljárás megindítása és szankciók kiszabása előtt mindig javasolt egy informális találkozó összehívása, és a probléma megvitatása formális kereteken kívül, mert legtöbbször ez is elegendő a helyzet kezelésére. A jelentésben szereplő személy gyakran a kapott tájékoztatás ellenére sincs azzal tisztában, hogy cselekménye sérti az iskola bántalmazás elleni szabályzatát. Továbbá, egyes esetekben a magatartás még nem éri el a szabályzat megsértésének szintjét, de annak *folytatódásával súlyosabb cselekménnyé* válhat. Ilyen esetben egy szakértővel történő személyes beszélgetés segíthet megállítani ezt a folyamatot. Jelen tanulmány támogatja ezt a megoldást, hiszen nem célja a tanuló azonnali megbüntetése és formalizált eljárásban való felelősségre vonása, hanem tanítani, nevelni igyekszik a diákokat.

Ugyanakkor, ezek az informális találkozók nem mindig vezetnek eredményre. Az anti-bullying szabályzatnak ezért meg kell határoznia komolyabb szankciókat a tanulóokra, alkalmazottakra egyaránt. Az oktatási szempontot figyelembe véve, az írásbeli figyelmeztetés vagy a képzéseken, workshopokon való kötelező jelenlét az első lépés a szankciók sorában. *Különböző témájú tréningek (például döntéshozatali workshop vagy resztoratív technikák megismerése)* segítséget nyújthatnak az érintetteknek, hogy elkerüljék a további felelősségre vonást. További szankcióként kiszabható, hogy a bán-

talmazó kötelező kutatást folytasson a bullying *témájában*, vagy írjon egy esszét a viselkedéséről, *annak okairól*.

Ha komolyabb beavatkozás szükséges, akkor a klasszikus fegyelmi eljárásban alkalmazott szankciók kiszabására is lehetőséget kell biztosítani, mint például a felfüggesztés és az eltanácsolás. E ponton lehetőség nyílik arra, hogy az anti-bullying szabályzatban rögzítse az iskola, hogy a helyzet további kezelése *már nem a bántalmazás elleni szabályzat keretei és eszköztára* révén, hanem fegyelmi eljárásban folytatódik. Mindazonáltal, e súlyosabb jogkövetkezményekre is utalhat a szabályzat, hiszen a bullying és cyberbullying magatartások már számtalan esetben vezettek rendkívül súlyos traumákhoz hazánkban,⁴¹ *több esetben, sajnos öngyilkossághoz is.*⁴²

Továbbá, az alkalmazottak jogellenes magatartásaira is megoldást kell nyújtania a szabályzatnak, azzal, hogy itt nem a bántalmazás elleni szabályzat feladata az alkalmazottak felelősségre vonása, mivel e területen már létezik egy kimunkált felelősségi rendszer. A pedagógusok és a munkavállalók jogellenes magatartását érintő jogszabályokra hivatkozás elegendőnek mutatkozik a szabályzatban.

Összefoglalóan, a fent említett szabályozási elemek a következők: (i) személyi és időbeli hatály; (ii) definíciók; (iii) a tiltott magatartások; (iv) a tiltott magatartásokra vonatkozó jelentéstétel és eljárás; (v) jogkövetkezmények, szankciók. Ezek a minta-szabályzat vagy a meglévő intézményi szabályzatok kiegészítésének a közös összetevői, amelyek együttesen képezik egy iskolai anti-bullying szabályzat legfontosabb alapköveit. E főbb pontok részleteinek kialakítását pedig az adott oktatási intézményre kell bízni, mert ez lehetővé teszi, hogy saját intézményükre szabják azt. Végül megfelelő képzésekkel, minta-szabályzatokkal és szakmai segítséggel szükséges támogatni őket, amelyet a fenntartónak kell folyamatosan biztosítania.

4. Konklúzió

Tekintettel a fent leírtakra, az offline és az online bántalmazás jelen van a magyar iskolák mindennapjaiban, ám az intézmények sokszor védtelenek, kiszolgáltatottak e harcban, a megfelelő ismeretanyag és jogi háttér nélkül. Átfogó, egységes, egész köznevelési rendszerre kiterjedő bántalmazás elleni program, valamint törvény hiányában rendkívül nehéz helyzetbe kerülnek az intézmények.

Ennek megkönnyítése érdekében született meg közös minimumként az alább meg-szövegezett bántalmazás elleni minta-szabályzat, amelynek célja, hogy támpontként szolgáljon egy anti-bullying szabályzat kialakításakor. Elkülönülve a magyarázatokban kifejtettektől, a javasolt minta-szabályzat szövege a Mellékletben olvasható.

⁴¹ Kulcsár i. m.

⁴² NF: Holtan találták a vajai tinit, aki iskolai zaklatások után tűnt el. Index, 2021. 09. 07. <https://tinyurl.com/4abnn5ae> ; Krupincza Mariann: Sümegi tragédia: miért fajulhatott idáig a helyzet, min mennek át a bántalmazottak és a bántalmazó tinédzserek? *Mandiner*, 2023. november 21. <https://tinyurl.com/29whz63r> ; Nagy Nikolett: Vonat gázolt el egy középiskolai diákot Sümegen. *Telex*, 2023. november 19. <https://tinyurl.com/2ctsmajj>

Melléklet

Minta-szabályzat az offline és online bántalmazás (bullying és cyberbullying) megelőzésére és kezelésére

Az Alaptörvényben és a 2011. évi CXCV. törvény a nemzeti köznevelésről című törvényben foglalt alapvető jogok és szabadságok tiszteletben tartása és védelme céljából, különös tekintettel a köznevelési intézményekben tanulók méltóságának, testi és szellemi fejlődésének védelmére, valamint nevelésük erkölcsi megerősítésére, a(z)(nevelési-oktatási intézmény neve) a következő szabályzatot fogadja el, elítélve a bántalmazás minden formáját, s minden módon elősegítve annak megelőzését, kezelését, valamint az elszenvedett sérelmek kompenzálását:

1. § (1) A szabályzat hatálya kiterjed az (nevelési-oktatási intézmény neve) valamennyi gyermekére, tanulójára, azok szüleire, törvényes képviselőire, valamint a nevelőtestület minden tagjára és a nevelési-oktatási intézmény minden alkalmazottjára, továbbá az említett személyekkel az intézményi működés keretében rendezett módon kapcsolatba került személyekre.

(2) A szabályzatot alkalmazni kell minden olyan helyszínre, rendezvényre, programra és egyéb eseményre, amely a nevelési-oktatási intézmény területén, tanítási időben zajlik. Továbbá a hatály kiterjed a nevelési-oktatási intézményhez kötődő egyéb programokra, amelyek az intézmény nevelési-oktatási feladatával állnak összefüggésben, így különösen osztálykirándulás, tanítási időn kívüli intézménylátogatás, színházi előadás, ifjúsági és kulturális rendezvényen részvétel, egyéb.

(3) Az elektronikus kommunikáció útján megvalósított minden, jelen szabályzatban meghatározott bántalmazás, amely az nevelési-oktatási intézmény területén, tanítási időben, vagy intézményhez kötődő programon akár intézményi tulajdonban, akár magántulajdonban álló készüléken történik, jelen szabályzat hatálya alá tartozik.

2. § E szabályzat alkalmazásában:

1. tanuló: a 20/2012. (VIII.31.) EMMI rendelet 1.§ b) alapján a nevelési-oktatási intézménybe felvett gyermek;

2. szülő: a 20/2012. (VIII.31.) EMMI rendelet 1.§ d) alapján a gyermekek, tanulók szülei, törvényes képviselői;

3. alkalmazott: a 20/2012. (VIII.31.) EMMI rendelet 1.§ c) alapján a pedagógus és a nevelő és oktató munkát közvetlenül segítő munkakörben alkalmazottak, az óraadók, valamint a nevelési-oktatási intézményben nem pedagógus-munkakört betöltők;

4. az intézményi működés keretében rendezett módon kapcsolatba került személyek köre: e körbe tartozhat különösen az önkéntes, a nevelési-oktatási működést segítő civil, gazdasági, egyházi és önkormányzati szervezetek munkatársa, képviselője;

5. elektronikus kommunikáció: az elektronikus kommunikáció magában foglalja, de nem kizárólag, a telefonhívásokat, szöveges üzeneteket (SMS), e-maileket, közösségi médiát és az azonnali üzeneteket az online térben;

6. offline bántalmazás (bullying): ismétlődő és károkozó magatartás bármely tanuló, szülő, alkalmazott vagy az intézményi működés keretében rendezett módon kapcsolatba került személy által, illetve annak sérelmére fizikai és/vagy lelki sérelem, tulajdonában bekövetkező kár okozása vagy arra irányuló szándék, amely erőegyensúly felborulást eredményez vagy tart fenn. Fizikai és/vagy lelki sérelem okozása különösen a bántalmazott nevetségessé tétele, megalázása, lealacsonyítása vagy dehumanizálása;

7. online bántalmazás (cyberbullying): a tanuló, szülő, alkalmazott vagy az intézményi működés keretében rendezett módon kapcsolatba került személy által, vagy sérelmére, elektronikus kommunikáció által elkövetett, szándékos, egyszeri vagy ismétlődő magatartás, amelynek célja az erőegyensúly felborítása, s megvalósítja az alábbiak bármelyikét:

- a) testi-lelki egészség sérelme vagy tulajdonban való károkozás, illetve ezek bekövetkezése a körülmények alapján valószínűsíthető; vagy
- b) ellenséges oktatási környezet létrehozása, vagy az iskola által nyújtott szolgáltatásoktól, hasznoktól és lehetőségektől való megfosztás;

8. megtorlás: megtorlásnak minősül az a magatartás, amely a bántalmazás elleni szabályzatban foglaltak megsértése miatt jelzéssel élő, eljárást indító vagy az eljárásban közreműködő tanulóval, szülővel, alkalmazottal vagy az intézményi működés keretében rendezett módon kapcsolatba került személlyel szemben ezzel összefüggésben sérelmet okoz, sérelem okozására irányuló magatartást tanúsít, vagy azzal fenyeget. Megtorlás különösen, ha a jelentéstétel tényleges bántalmazó magatartás hiányában valósul meg.

3. § (1) Bántalmazást követ el, aki jelen szabályzat 2.§ 6. vagy 7. pontjában foglalt magatartást tanúsít a nevelési-oktatási intézmény területén, tanítási időben, vagy intézményhez kötődő programon.

(2) Megtorlást követ el, aki jelen szabályzat 2.§ 8. pontjában foglalt magatartást tanúsít az nevelési-oktatási intézmény területén, tanítási időben, vagy intézményhez kötődő programon.

4. § (1) (A nevelési-oktatási intézmény neve) elkötelezett a bántalmazás elleni küzdelemben. Ennek elsődleges területe a prevenció. Ezt szolgálja, hogy jelen szabályzatot elérhetővé kell tenni az elfogadását (módosítását) követő 24 órán belül elektronikus úton a nevelési-oktatási intézmény honlapján, valamint az intézmény több pontján, nyomtatott formában.

(2) A(z) (nevelési-oktatási intézmény neve) közösségének minden tagja köteles megismerni a bántalmazás elleni szabályzatban foglaltakat, és köteles az abban foglaltakat betartani. Ennek következtében minden frissen beiratkozó tanulóknak és szüleinek kötelező a bántalmazás elleni szabályzat megismertetésére irányuló foglalkozáson részt venni.

(3) Minden alkalmazott köteles a bántalmazás elleni szabályzat megismerésére irányuló képzésen részt venni. A szaktudás hatályos és naprakész ismerete érdekében minden alkalmazott köteles ezen képzést tanévenként megismételni.

5. § Bántalmazás esetén bárki, aki tudomással bír a bántalmazásról, jelentéssel élhet. A bántalmazás jelentése történhet személyesen az arra kijelölt személy(ek)nél, anonim módon az erre a célra kihelyezett jelentéstételi ládában, online, névvel ellátott jelentéstételi űrlap kitöltésével, vagy online, anonim módon kitöltött jelentéstételi űrlap útján. A jelentéstétel során, offline és online egyaránt, benyújthatóak a bántalmazást bizonyító dokumentumok, képernyőmentések, videófelvetelek és egyéb eszközök.

6. § A bántalmazásról érkezett jelentés alapján, az ügy megvizsgálása és arról szakértői vélemény készítése a kijelölt személy(ek) feladata. E vélemény elkészítése során a kijelölt személy(ek) informális és formalizált találkozókat szervezhet(nek), szabadon értekelhetik a bizonyítékokat és rendelkezésre álló információkat. A szakértői véleményben meg kell állapítani, hogy megvalósult-e bántalmazás vagy megtorlás. Amennyiben igen, úgy cselekvési tervet kell készíteni.

7. § (1) A szakértői vélemény és a cselekvési terv elkészítését követően az esetkezelésért felelős személyhez vagy szervezeti egységhez (csapat) kell utalni. Az esetkezelésért felelős szervezeti egység tagjai között legalább egy fő a szakértői véleményt elkészítő személyek közül kerül ki.

(2) A cselekvési terv fő célja a bántalmazó magatartás abbahagyása, megszüntetése, a bántalmazó egyidejű felelősségvállalásával és jóvátétel iránti szándékának egyértelmű kinyilvánításával.

(3) A cselekvési terv végrehajtása, szükség esetén módosítása, az esetkezelésért felelős személy vagy szervezeti egység (csapat) feladata. A cselekvési terv végrehajtását két alkalommal kell megkísérelni. Amennyiben a két alkalom megkísérlését követően a bántalmazás nem szűnt meg, úgy az esetet a döntéshozatali szerv elé kell utalni a szakértői vélemény és a cselekvési terv sikertelen végrehajtását igazoló dokumentáció egyidejű megküldésével.

(4) Az esetkezelésért felelős személy vagy szervezeti egység (csapat) a szakértői vélemény és a (3) bekezdésben foglalt dokumentáció megküldésével egyidejűleg köteles javaslatot tenni a szankció kiszabására és annak formájára. A döntéshozó szerv a javaslat figyelembevételével dönt a szankciók alkalmazásáról, de a bántalmazás tényét vagy annak hiányát megállapító szakértői véleményt nem bírálhatja felül.

8. § (1) A döntéshozatali szerv a Bántalmazási Elleni Tanács (a továbbiakban: BET). A BET olyan testület, amelynek tagja az iskola igazgatója, igazgatóhelyettese/í, a szakértői javaslatot elkészítő személy (szervezeti egység esetén, annak képviselőjében egy fő), az esetkezelésért felelős személy (szervezeti egység esetén, annak képviselőjében egy fő), valamint a tanulók által kijelölt tanuló.

(2) A BET eljárása nem nyilvános, annak ülésein kizárólag az érintett felek és a BET által meghívott személyek vehetnek részt. Döntését elsődlegesen a szakértői véleményre és az esetkezelés során készült dokumentumokra alapozza, de az érintetteket meghallgathatja, a rendelkezésre álló bizonyítékokat, iratokat megismerheti. Döntéseit egyszerű szótöbbséggel hozza. Határozatképességéhez legalább a tagok több mint felének jelenléte szükséges.

(3) A törvényes képviselően túl más (például jogi, szakszervezeti, érdekképviseleti, munkáltatói) képviselő *részvétele az eljárásban kizárt.*

9. § (1) A BET döntése alapján kiszabható szankciók tanulók, szülők és alkalmazottak esetében a *következők:*

- a) írásbeli figyelmeztetés;
- b) bántalmazás elleni képzésen való kötelező részvétel;
- c) resztoratív technikákat ismertető képzésen való kötelező részvétel;
- d) öt alkalommal való kötelező részvétel iskolapszichológusi ülésen, vagy
- e) fegyelmi eljárás megindítása.

(2) Alkalmazottak esetében a mindenkori munkajogi törvény által szabályozott eljárások és azok jogkövetkezményeinek alkalmazására tehet javaslatot a BET.

10. § A BET döntése ellen jogorvoslati eljárás indítható a mindenkori köznevelési törvényben meghatározottak szerint.

11. § Jelen szabályzatot a (nevelési-oktatási intézmény) (dátum) napon elfogadta és hatályba is lép. Jelen szabályzat visszavonásig, de legkésőbb _____ évig hatályos. Az alkalmazási idő vége előtt legkésőbb 90 nappal el kell végezni a végrehajtásra vonatkozó felülvizsgálatot (a hatékonyság, visszatartó erő, ismertség, elégedettség szempontjából), és annak eredménye alapján a BET javaslatot tesz a szabályzat megfelelő módosítására vagy változatlan szabályok szerinti további hatályba léptetésére (hatályba tartási javaslat).

A MESTERSÉGES INTELLIGENCIA ALAPÚ TARTALOMSZŰRŐ ALGORITMUSOK TÁRSADALMI HATÁSAI, KÜLÖNÖS TEKINTETTEL A KÖZÖSSÉGIMÉDIA-PLATFORMOKRA

Üveges István

1. Bevezetés

Napjainkra a közösségekben zajló kommunikáció elsődleges tere egyértelműen az online világ lett. A különböző online platformok, közösségimédia-felületek, videómegosztó oldalak stb. az információáramlás domináns médiumát képezik. Az ilyen, gyakran globálisan működő kommunikációs terek azonban fokozottan ki vannak téve a szélsőséges ideológiák, szerzői jogi szempontból aggályos, esetleg más szempontból jogsértő, gyakran felkavaró tartalmak ellenőrizetlen terjedése okozta kockázatnak.

A probléma kezelésére bevett gyakorlat, hogy az ilyen platformokra feltöltött tartalmak ellenőrzésére esnek át, akár automatikusan, közvetlenül a feltöltésük után, akár később, ha egy felhasználó problémásként bejelenti az adott tartalmat. Ennek következménye lehet többféle szankció is, a tartalom törlésétől a felhasználó kizárásán át egészen jogi eljárás megindításáig. Ezt a fajta ellenőrzési mechanizmust szokás összefoglalóan ‘moderálásként’ hivatkozni.

Azáltal, hogy például a közösségimédia-platformok mögött álló cégek mind saját megoldásokat alkalmaznak a felületeiken felbukkanó tartalmak moderálására, valójában olyan hatalomra tesznek szert, amellyel korábban csak államok és egyházak rendelkezhetek.¹ Vékony a határvonal az intézményesített cenzúra, és a felhasználók érdekeit szem előtt tartó tartalomszűrés között. A kettő összemosódása különösen gyakran kerül szóba annak kapcsán, hogy a legnagyobb közösségimédia szolgáltatók mind jobban támaszkodnak akár mesterséges intelligenciát is alkalmazó, de mindenképpen teljesen automatizált rendszerekre a moderálási folyamat során.

* Tudományos segédmunkatárs, ELTE Társadalomtudományi Kutatóközpont.

ORCID: <https://orcid.org/0000-0001-5897-9379>

¹ Gosztonyi Gergely: *Cenzúra Arisztoteléstől a Facebookig*. Budapest, Gondolat, 2022.



Az ilyen rendszerek működése során az emberi tényező háttérbe szorul, hiszen a döntéseket sok esetben kizárólag algoritmusok hozzák meg. Ez nem pusztán az ilyen rendszerek pontatlansága miatt lehet problémás, de például gépi tanulás alkalmazása esetében az azok tanítóadataiban jelenlévő torzítások (*bias*) is kockázati faktort jelentenek.

A jelen tanulmány célja éppen ezért, hogy feltárja az automatizált/algorithmikus moderálás elterjedése mögött álló tényezőket, felhívja a figyelmet azok kockázataira és hatásaira a társadalom egészére nézve, valamint, hogy egyfajta helyzetképet adjon az online kommunikációs terek körül kialakult viták aktuális állapotáról, néhány lehetséges megoldást is felvázolva. Ehhez elsőként a tanulmány áttekinti a jogi helyzetet, amely az Egyesült Államokban lehetőséget teremtett az online kommunikáció ma ismert formáinak elterjedésére. Ezt követően ismerteti, hogy milyen (nyilvánosan megismerhető) módszereket alkalmaznak az érintett platformok a náluk megjelenő tartalmak megfelelőségének biztosítására. Végül felvázol néhányat azon problémás pontok közül, ahol ezek a módszerek kihatnak az egyén szólásszabadságára, vagy éppen az egyenlő bánásmódhoz való jogára.

2. A szólásszabadság védőbástyája – és Achilles-sarka

Az online szólásszabadság kérdését a közösségi médiában alapvetően a *Communications Decency Act* (CDA)² 230-as szakasza határozza meg az USA esetében. Ez alapvető fontosságú, hiszen a legtöbb ma ismert közösségimédia-platform innen származik. A CDA megszületésének háttéréről több átfogó elemzés³ is napvilágot látott.⁴ Ezek alapján rekonstruálható a törvénybe, főként annak leginkább vitatott 230-as szakaszába foglaltak háttére.

Az online tartalommoderálás szükségességét érintő viták talán egyik első sarkalatos pontját egy 1995-ös USA-beli per, a Stratton Oakmont v. Prodigy Services ügy jelentheti.⁵ A per első résztvevője, a felperes, a Stratton Oakmont Inc. volt, egy Long Island-i székhelyű brókerház, amelyet Jordan Belfort alapított 1989-ben. A vállalat elsősorban alacsony áron beszerezhető, tőzsdén kívüli „centes részvényekkel” (*penny stock*) kereskedett, és az 1990-es években már a legnagyobb tőzsdén kívüli részvények kereskedésére szakosodott vállalat volt az USA-ban.⁶

² 47 U.S.C. § 230 (1996).

³ Például David S. Ardia: Free Speech Savior or Shield for Scoundrels: An Empirical Study of Intermediary Immunity under Section 230 of the Communications Decency Act. *Loyola of Los Angeles Law Review*, Vol. 43. (2010) 373–506.; Benjamin Volpe: From Innovation to Abuse: Does the Internet Still Need Section 230 Immunity? *Catholic University Law Review*, Vol. 68. (2019) 597–624.; vagy éppen Jeff Kosseff: A User’s Guide to Section 230, and a Legislator’s Guide to Amending It (or Not). *Berkeley Technology Law Journal*, Vol 37., No. 2. (2022) 757–802. <https://doi.org/10.15779/Z38VT1GQ97>

⁴ Jeff Kosseff: *The Twenty-Six Words That Created the Internet*. Ithaca, NY, Cornell University Press, 2019.

⁵ Stratton Oakmont, Inc. v. Prodigy Services Co., 1995 WL 323710 (N.Y. Sup. Ct. May 24, 1995).

⁶ Matthew Partridge: Great frauds in history: Jordan Belfort and Stratton Oakmont. *MoneyWeek*, 2019. augusztus 7. <https://tinyurl.com/3tv7wcy4>

Az alperes egy korai internetes tartalomszolgáltató, a Prodigy Services Co.⁷ volt. A cég fő profilja információs szolgáltatások nyújtása volt a feliratkozók részére, például „hirdetőtáblák” (*bulletin boards*) segítségével. Ezekon harmadik felek információt cserélhettek egymással, nagyjából a mai értelemben vett internetes fórumokhoz hasonló módon. A cég olyan online szolgáltatóként hirdette magát, mint aki a pusztán közzétételén túl szerkesztői ellenőrzést is gyakorol az általa üzemeltetett online platformokon megjelenő tartalmak felett. Geoffrey Moore, a Prodigy piaci programokért és kommunikációért felelős igazgatója a pert megelőzően több alkalommal is sajtónyilvánosság előtt hangoztatta, hogy ezáltal ők sokkal inkább a nyomtatott sajtóhoz hasonlítanak, semmint egy ellenőrizetlen online felülethez.⁸

A cég egyik igen népszerű fóruma a „Money Talk” volt, ahol a résztvevők pénzügyi kérdéseket vitathattak meg egymással, és amelyet mai szóhasználattal élve ‘moderátorok’ is felügyeltek. 1994-ben a fórumon aztán egy olyan bejegyzés jelent meg, amely a Stratton Oakmont-ot bűncselekményekkel és csalással vádolta. A Stratton Oakmont erre válaszul beperelte a Prodigy-t rágalmazásért. A per során a kulcskérdés az volt, hogy a Prodigy a megjelent információk kapcsán ‘kiadónak’ (*publisher*) vagy csak ‘terjesztőnek’ (*distributor*) minősült-e. A kérdés azért bírt különös relevanciával, hiszen az USA-beli esetjog hagyományosan erősen elkülöníti egymástól az e két szerepkörrel járó felelősséget.

E szerint a megkülönböztetés szerint egy *kiadótól* elvárható, hogy tudatában legyen az általa közzétett anyag tartalmának, jellegének, valóságának vagy valóságtartalmának, így tehát felelősségre vonható az általa közzétett jogellenes tartalmakért. Ezzel szemben egy *terjesztőnek* valószínűleg nincs tudomása az általa elérhetővé tett anyagok tartalmáról, így pedig mentesülhet az azzal összefüggő jogi felelősség alól is.⁹

Ahogy például a Doe v. America Online esetben, a Zeran v. America Online esetből idézve az Egyesült Államok Negyedik Kerületi Fellebbviteli Bírósága ekkor megállapította, a rágalmazási perekben alapvető elem a ‘publikálás’ (~kiadás) fogalma, ami azt jelenti, hogy a közölt anyagok tartalmáért csak az vonható felelősségre, aki azokat publikálja. A publikálás itt nem csak az információk közzétételére vonatkozik, hanem a rágalmozó állítás gondatlan közlésére vagy a mások által elsőként közölt állítás eltávolításának elmulasztására is.¹⁰

A Prodigy szempontjából a döntés végül hátrányosan alakult, hiszen a bíróság a cég fentebb említett nyilatkozatai, és az alkalmazott moderálási elvei miatt *kiadóként* tekintett rá, ennek megfelelően pedig felelősnek is tekintette az általa gondozott oldalakon megjelent állításokért. Ez egyébként részben szembement egy korábbi, 1991-es

⁷ Prodigy Communications Corporation History. FundingUniverse, é.n. <https://tinyurl.com/mr3zs69v>

⁸ Stratton Oakmont, Inc. v. Prodigy Services Co., No. 31063/94, 1995 WL 323710 (N.Y. Sup. Ct. May 24, 1995), opinion by Justice Stuart L. Aín.

⁹ Richard J. Hunter Jr. – Hector R. Lozada – John H. Shannon: Distributor vs. Publisher vs. Provider: That Is the High-Tech Question: But is an Extension of Liability the Answer? *International Journal of Education and Social Science*, Vol. 8., No. 1. (2021) 28–36.

¹⁰ Doe v. America Online, Inc. (2001). 783 So. 2d 1010 (Florida Supreme Court); Zeran v. America Online, Inc. (1997). 129 F. 3d 327 (United States Court of Appeals, 4th Circuit).

döntéssel, amikor is a bíróság a CompuServe-et, mint egy oldal üzemeltetőjét csak mint ‘terjesztőt’ ismerte el.¹¹

Az eset kapcsán szárnyra kaptak olyan álláspontok, hogy a Stratton Oakmont kapcsán hozott döntés inkább visszalépés a korábbi gyakorlathoz képest, hiszen ezen állásfoglalásával a bíróság arra biztatja az internetes platformok üzemeltetőit, hogy semmilyen tartalommoderálást ne végezzenek, ezzel elhárítva mindennemű felelősséget az USA-beli jog alapján. Voltaképpen ennek az ellentmondásnak a feloldása lett az, amely később a manapság is híres/hírheft CDA 230. szakaszában öltött testet.

A *Telecommunications Act of 1996*¹² részeként elfogadott *Communications Decency Act* (CDA)¹³ biztosítja a platformok számára a jogi védelmet a felhasználói tartalommal kapcsolatos felelősség alól, miközben lehetővé teszi számukra a tartalom moderálását is. A CDA 230. szakasza kimondja, hogy az online szolgáltatók nem felelősek a felhasználók által közzétett tartalomért,¹⁴ és biztosítja számukra az immunitást a moderációs döntéseikért, feltéve, hogy jóhiszeműen jártak el egy-egy károsnak ítélt tartalom eltávolítása során.¹⁵ E védelem azonban nem terjed ki minden esetre: a szerzői jogi igényeket a *Digital Millennium Copyright Act* (DMCA), a szexuális szolgáltatásokkal kapcsolatos tartalmakat pedig a FOSTA/SESTA törvénycsomag kifejezetten kivonja a 230. § oltalma alól.

Egészen pontosan a 230 (c)(1) bekezdése a következők szerint fogalmaz:

Egyetlen interaktív számítógépes szolgáltatás szolgáltatója vagy felhasználója sem tekinthető egy másik információs tartalomszolgáltató által nyújtott információ kiadójának (*publisher*) vagy előadójának.

Az érem másik oldalát közvetlenül ezután a (2) bekezdés tárgyalja:

Egyetlen szolgáltató vagy interaktív számítógépes szolgáltatás felhasználója sem vonható felelősségre az alábbiak miatt:

- (A) bármely, a szolgáltató vagy felhasználó által obszcénnek, kéjesnek, bujának, mocskosnak, túlzottan erőszakosnak, zaklatónak vagy más módon kifogásolhatónak tartott anyaghoz való *hozzáférés* vagy *ilyen tartalom elérhetőségének korlátozása érdekében jóhiszeműen önkéntesen tett intézkedésért, függetlenül attól, hogy az ilyen anyag alkotmányos védelem alatt áll-e vagy sem*; vagy
- (B) bármely olyan intézkedés, amely lehetővé teszi vagy elérhetővé teszi az információs tartalomszolgáltatók, vagy mások számára az (1) bekezdés-

¹¹ *Cubby, Inc. v. CompuServe Inc.*, 776 F. Supp. 135 (S.D.N.Y. 1991).

¹² *Telecommunications Act of 1996*, Pub. L. No. 104-104, 110 Stat. 56 (1996).

¹³ Ruth Ann Strickland: *Telecommunications Act of 1996* (1996). *Free Speech Center*, January 1, 2009. <https://firstamendment.mtsu.edu/article/telecommunications-act-of-1996/>

¹⁴ Anthony Ciolli: *Chilling Effects: The Communications Decency Act and the Online Marketplace of Ideas*. *University of Miami Law Review*, Vol. 63. (2008) 137. <https://doi.org/10.2139/ssrn.1101910>

¹⁵ Richard J. (2021) i. m. 30–32.

ben leírt anyagokhoz való hozzáférés korlátozására szolgáló technikai eszközöket.

Ahogy azt a *Doe v. GTE Corp.* ügyben¹⁶ a bíróság kifejtette, a 230. szakasz (c) (1) és (2) bekezdés közötti különbség az, hogy az első „megakadályozza a polgári jogi felelősségre vonást, ha a webtárhelyek és más internetszolgáltatók (ISP-k) tartózkodnak az oldalukon található információk szűrésétől vagy cenzúrázásától”, míg az utóbbi biztosítja, hogy a „sértő anyagokat kiszűrő szolgáltató ne legyen felelősségre vonható a cenzúrázott ügyfél által”.¹⁷

Egy másik eltérés, hogy a 230. (c)(2) szakasz védelme csak a „jóhiszeműen tett” intézkedésekre vonatkozik. A „jóhiszeműség követelménye” nem létezik a 230. (c)(1) szakasz tágabb értelemben vett rendelkezéseiben. A bíróságok kifejtették, hogy ez a követelmény a kifogásolható tartalmakat eltávolító vagy az azokat tévesen el nem távolító online szolgáltatások védelmére szolgál, anélkül, hogy azokat is védelemben részesítené, akik versenyellenes vagy más rosszindulatú szándékból (visszaélésszerűen) távolítanak el tartalmakat.¹⁸

A fentiek miatt végül a CDA 230. szakasza lett az internetes szólásszabadság védőbástyája, és egyben Achilles-sarka is. Azáltal, hogy a rendelkezés megvédi az online szolgáltatásokat a harmadik fél által készített tartalmakért való felelősségre vonástól, egyben meg is nyitotta az utat a különböző, ma már széles körben elterjedt üzleti modellek előtt, amelyek a felhasználók által generált tartalmakra támaszkodnak, átalakítva az ekkor még csak formálódóban levő online gazdaságot is. Lényegében ez tette lehetővé, hogy olyan tudásmegosztó oldalak, mint a Wikipédia, szabadon gyakorolhassanak tartalommoderálást a felhasználói bejegyzések felett. De ugyanígy ez teremtett lehetőséget például a kis webshopok üzemeltetői számára, hogy a forgalmazott termékeikről szóló vásárlói véleményeket tehessenek közzé a gyártók retorzióitól való félelem nélkül.

3. Az automatikus moderáláshoz vezető út

A CDA 230. szakasza által nyújtott szabadság tehát hatalmas üzleti lehetőségeket nyitott, egyúttal azonban számos probléma kialakulásának is megágyazott. Az egyik sarkalatos pont például, amely körül a viták az USA-ban újra és újra fellobbannak, a terrorizmussal kapcsolatos, pontosabban a terrorista szervezetek által közzétett on-

¹⁶ *Doe v. GTE Corp.*, 347 F.3d 655 (7th Cir. 2003).

¹⁷ A terület hazai irodalmával kapcsolatban lásd különösen: Imre Melinda: Az internet-szolgáltatók felelősségének szabályozása a szerzői jogot sértő tartalmak tekintetében – Az amerikai, a közösségi és a magyar szabályozás bemutatása. *Iustum Aequum Salutare*, 2006/1–2. 213–227.; Muraközi Gergely: A szerzői jog és az internet – Az internet technikai megvalósítása a szerzői jog tükrében. *Jogi Fórum*, én. [https://www.jogiforum.hu/files/publikaciok/drMurakozi-A_szerzoi_jog_es_az_internet\(jf\).pdf](https://www.jogiforum.hu/files/publikaciok/drMurakozi-A_szerzoi_jog_es_az_internet(jf).pdf); Gosztonyi Gergely: A platformszolgáltatók felelősségének új szabályozása az európai uniós digitális szolgáltatásokról szóló rendelet alapján. *Pro Futuro*, 2023/3. 3–26.

¹⁸ Ash Johnson – Daniel Castro: Overview of Section 230: What It Is, Why It Was Created, and What It Has Achieved. *Information Technology & Innovation Foundation*, 2021. február 22. <https://tinyurl.com/mr2w953n>

line tartalmak kezelése. A leggyakoribb téma ennek kapcsán a szólásszabadság és a nemzetbiztonság közötti egyensúly kérdése, amely napjainkban is élénk jogi és társadalmi viták tárgyát képezi.¹⁹ Az Egyesült Államok alkotmánya, különösen az Első Alkotmánykiegészítés²⁰ garantálja a szólásszabadságot, amely magában foglalja az online tartalmak szabad megosztását is. Ugyanakkor, ha ezek a tartalmak terrorizmust támogató vagy erőszakra buzdító üzeneteket tartalmaznak, a hatóságok és a közvélemény egy része azt érzi, hogy a nemzetbiztonsági kockázatok miatt szükséges ezen tartalmak kiemelten hatékony korlátozása vagy eltávolítása. A fő kérdés ennek kapcsán a moderálás hatékonyságának állami kikényszeríthetősége alapvetően magántulajdonban levő vállalatokra vonatkozóan.

A helyzetet bonyolítja, hogy az Egyesült Államokban a kezdeti időszakban a kormányzat az újra és újra kieleesedő viták ellenére elsősorban nem közvetlen jogszabályi úton kívánta rendezni az olyan kérdéseket, amelyek hosszabb távon problémásnak bizonyultak.²¹ Ennek hátterében nem kizárólag a közvélemény megosztottsága állt, hanem az is, hogy a platformok felhasználói bázisa, valamint az általuk kezelt tartalmak mennyisége ekkor még lényegesen kisebb volt, mint napjainkban. A jogalkotók ezért elsősorban kevésbé formális, együttműködésen alapuló megoldásokat részesítettek előnyben, például köz- és magánszféra közötti partnerségeket, figyelemfelkeltő kampányokat, egyetértési megállapodásokat, a rendelkezésre álló szakértelem megosztását, valamint a felelőségek elhatárolását az amerikai kormány és a nagy közösségimédia-vállalatok között. Ezt az időszakot a szakirodalom gyakran „társszabályozásként” (*co-regulation*) jellemzi, amely a szabályozói szándék és a piaci szereplők önszabályozása közötti egyensúly kialakítására törekedett.²²

Persze emellett az sem elhanyagolható szempont, hogy például a nagy közösségimédia-platformok mindegyike a tőzsdén is bejegyzett vállalat. A céggel kapcsolatos negatív hírek sok esetben direktben megmutatkoznak a részvényárfolyamok csökkenésében is, ami pedig közvetetten, a befektetőkön keresztül erős nyomást gyakorol az érintett cégek döntéshozóira is.²³ Mindezek mellé adódik még a szociális normák szerepe, vagyis, hogy sok esetben maga a szolgáltatást használó közösség az, aki értékrendjét az online platformokon megjelenő tartalmakban is szeretné viszont látni, illetve ellenkező esetben ezt az értéktranszfert az üzemeltető irányában kikényszeríteni.

Le kell azt is szögezni, hogy a közösségimédia-szolgáltatók is felelősek az olyan tartalmak korlátozásáért, amelyek törvénybe ütközőek, például éppen az *Anti-Terrorism*

¹⁹ Gosztonyi Gergely – Lendvai Gergely Ferenc: Twitter kontra Taamneh és Gonzalez kontra Google, avagy ki a felelős az online platformokra feltöltött tartalomért? *Magyar Jog*, 2023/10. 587–593.

²⁰ U.S. Const. amend. I.

²¹ Alexander Tthesis: Social Media Accountability for Terrorist Propaganda. *Fordham Law Review*, Vol. 86. (2017) 605–631.

²² Koltay András – Nyakas Levente (szerk.): *Magyar és európai médiajog*. Budapest, Wolters Kluwer, 2017.; Lawrence Lessig: What Things Regulate Speech. In: *Code: And Other Laws of Cyberspace*. New York, Basic Books, 1999. 86–88.

²³ Nicole Softness: Terrorist Communications: Are Facebook, Twitter, and Google Responsible for the Islamic State's Actions? *SIPA Journal of International Affairs*, Vol. 70. No. 1. (2017) 201–215.

*Act*²⁴ hatálya alá tartozó esetekben, vagy szerzői jogot sértő tartalmak esetében. Részen emiatt, részben pedig a törvényileg nem szabályozott, de a társadalom értékrendjét esetlegesen sértő esetek szabályozása miatt a legtöbb platform a felhasználási feltételeiben (*Terms of Service*, a továbbiakban: ToS) rendelkezik arról, hogy milyen tartalmakat töröl, vagy mely esetekben szankcionálja a tartalom közzétevőjét. Ez a szankció a legtöbb esetben az érintett tartalom eltávolításán keresztül valósul meg, amelyet a szolgáltatók – például a folyamatos társadalmi nyomásra reagálva – igyekeznek hatékonyan, és legfőképpen, gyorsan végrehajtani. Ugyanakkor nem kizárólag a tartalom törlése jelenthet szankciót: előfordulhat a felhasználói fiók ideiglenes vagy végleges felfüggesztése, a tartalom láthatóságának korlátozása (pl. *shadow banning*), földrajzi alapú korlátozása (*geoblocking*), vagy az elérés algoritmikus háttérbe szorítása is. Ezek a módszerek gyakran kevésbé átláthatók, és a felhasználók számára nehezebben követhetők nyomon, ami külön kérdéseket vet fel az elszámoltathatóság és a transzparencia kapcsán.

Az elmúlt éveket jellemzi, hogy az online tartalmakra vonatkozó szabályozások fokozatosan egyre inkább szigorodtak. A ma legnépszerűbb közösségimédia-platformok mögött álló szolgáltatók többsége az USA-ban van bejegyezve, ugyanakkor, tekintettel arra, hogy tevékenységük a világ számos országára kiterjed, így érvényesek rájuk például az Európai Unió területén annak vonatkozó jogszabályai is. Ezek közül talán a legfontosabb az EU digitális szolgáltatásokról szóló jogszabálya (*Digital Services Act*, DSA)²⁵, amely 2024. február 17-én lépett hatályba teljeskörűen. Ennek értelmében új kötelezettségek hárulnak azokra az online platformokra, amelyeknek az EU-ban is vannak felhasználói, azzal a céllal, hogy a felhasználók, illetve a felhasználók jogainak védelme hatékonyabbá válhasson.²⁶ A DSA rendelkezései több tucatnyi új előírást tartalmaznak, ám a tartalommoderálás és felhasználói jogvédelem szempontjából négy olyan pillér emelkedik ki, amely közvetlenül alakítja a platformok és a felhasználók mindennapi viszonyát.

- „*Notice-flag*” rendszer: ez az első, gyakorlatban már rövid távon érezhető újítás, mert kifejezetten a moderációs folyamat elejét – a problémás tartalom bejelentését – szabványosítja, így a felhasználók ténylegesen élhetnek a jogsértő bejegyzések gyors jelzésének lehetőségével.
- Kiskorúaknak célzott hirdetések tilalma: tiltja a kiskorúak profilalkotáson vagy személyes adataikon alapuló hirdetésekkel való megcélzását (pl. mikrotargetálás²⁷). A DSA ezzel a legvédtelenebb csoportot célozza, ahol a profilalkotásból fakadó károk – például egészség vagy testképtorzító üzenetek – a legnagyobbak.

²⁴ Anti-Terrorism Act (ATA), 18 U.S.C. § 2333.

²⁵ Az Európai Parlament és a Tanács 2022/2065/EU rendelete (2022. október 19.) a digitális szolgáltatásokról és a 2000/31/EK irányelv módosításáról (digitális szolgáltatásokról szóló rendelet), PE/30/2022/REV/1, HL L 277., 2022.10.27., 1–102. o.

²⁶ New rules to protect your rights and activity online in the EU. European Commission, 2024. február 16. <https://tinyurl.com/yc44sre3>

²⁷ Anja Prummer: Micro-targeting and polarization. *Journal of Public Economics*, Vol. 188. (2020) 104210. <https://doi.org/10.1016/j.jpubeco.2020.104210>

- Reklám-átláthatósági kötelezettség: a platform-gazdaság üzleti modellje hirdetés-vezérelt, így az információs aszimmetria csökkentése (miért pont ezt látom, ki fizetett érte?) kulcslépés a felhasználói autonómia megerősítéséhez.
- Érzékeny adatokon alapuló célzás teljes tiltása: teljes egészében betiltja az olyan hirdetéseket, amelyek érzékeny adatok, például politikai vagy vallási meggyőződés, szexuális preferenciák stb. alapján célozzák meg felhasználók egy csoportját.

A DSA hatálybalépése után a szabályok, amelyek 2023 óta már eddig is vonatkoztak néhány kiugróan nagy platformra és keresőmotorra, immár minden platformra és tárhelyszolgáltatásra érvényesek. Nem meglepő tehát, hogy különösen a közösségimédia szolgáltatók esetében, amelyek profitjuk jelentős részét adatkereskedelemből és hirdetések értékesítéséből nyerik, kiemelten fontossá vált ezek betartása, illetve betartatása.²⁸

Ahhoz, hogy a jogsértő tartalmakat gyorsan és hatékonyan szűrhessek ki, a platformok hagyományosan moderátorokat alkalmaznak, vagyis olyan embereket, akik feladata az oldalakra kikerülő tartalmak ellenőrzése, a beérkező bejelentések felülvizsgálata, illetőleg a szükséges szankciók kiszabása volt. Különböző algoritmikus megoldások is léteztek már egészen a kezdetektől, de a domináns irányvonal sokáig a *human-in-the-loop* megközelítés volt. Ez azt jelentette, hogy a legtöbb esetben a döntések legalább minimális emberi felülvizsgálat után születhettek csak meg.²⁹ Az utóbbi években azonban az automatizált előszűrés került túlsúlyba, amit a *Digital Services Act* (DSA) is elismer, ugyanakkor egyértelművé tesz, hogy az eltávolítási vagy korlátozási döntések nem alapulhatnak kizárólag automatizált megoldásokra. A szabályozás további újdonsága az átláthatóság követelménye. Ennek értelmében a legnagyobb szolgáltatóknak éves jelentésekben kell feltárniuk moderációs gyakorlatukat, és kutatók számára hozzáférést biztosítaniuk a releváns adatokhoz. Bár az első egységes DSA-jelentések csak 2025 tavaszán válnak elérhetővé, jelenleg a közölt információk inkább összesítő jellegűek, így a konkrét döntési logikák továbbra is nagyrészt rejtve maradnak.

A moderálási eljárást érő kiritkák már az automatizált módszerek bevezetése előtt is jellemzőek voltak. A New York University Stern, Centre for Business and Human Rights 2020 júniusában elkészített jelentése például a Facebookra, mint a legnagyobb közösségimédia platformra vonatkozóan az alkalmazott moderátorok kiszervezését, és a moderátorok számának elégtelenségét kritizálja.³⁰ A moderálás mikéntje aztán a COVID

²⁸ Üveges István: A mesterséges intelligencia közösségi médiában történő alkalmazásának társadalmi és politikai következményei. In: Kovács Zoltán (szerk.): *A mesterséges intelligencia és egyéb felforgató technológiák hatásainak átfogó vizsgálata*. Budapest, Katonai Nemzetbiztonsági Szolgálat, 2023. 301–327.

²⁹ Tom Lymn – Jessica Bancroft: The use of algorithms in the content moderation process. *Responsible Technology Adoption Unit Blog*, 2021. augusztus 5. <https://tinyurl.com/yt9x75a6>

³⁰ Paul M. Barrett: Who Moderates the Social Media Giants? A Call to End Outsourcing. NYU Stern Center for Business and Human Rights. *NYU STERN*, 2020 június. <https://tinyurl.com/4t2zuwfs>

hatására drasztikusan megváltozott.³¹ A pandémia alatt a közösségimédia platformok jelentős mértékben támaszkodtak algoritmikus tartalom moderációra, mivel a járvány következtében az emberi moderátorok munkája erősen korlátozottá vált. Az automatizált rendszerek bevezetése ekkoriban igencsak vegyes eredményeket hozott. Noha hatékonyak voltak a nagy mennyiségű tartalom kezelésében, az olyan problémákkal, mint a kontextus megértése és a különböző nyelveken való moderáció nehézségei, nem tudtak hatékonyan megbirkózni. A dezinformáció elleni küzdelem során a platformok gyakran szembesültek a politikai és tudományos viták bonyolultságával is, amely sokszor jelentősen megnehezíti a moderációs döntéseket. A tapasztalatok egyértelműen rávilágítottak az algoritmikus rendszerek korlátaira és az emberi felügyelet fontosságára.

A tanulmány eddigi részében a moderálás fogalma a felhasználói tartalmak feletti döntéshozatalra korlátozódott. Egyes szerzők meghatározásában azonban a *tartalommoderáció* ennél kiterjesztőbb értelmezést nyer, olyan irányítási mechanizmusok gyűjtőnevéként, amelyek strukturálják a közösségben való részvételt az együttműködés megkönnyítése és a visszaélések megelőzése érdekében.³² Fontos, hogy ebben az értelmezésben tehát a moderációs tevékenység részei természetesen a tartalom eltávolításáért felelős adminisztrátorok vagy moderátorok is, de ugyanígy beletartoznak azok a tervezési döntések is, amelyek alapjaiban megszervezik és befolyásolják, hogy egy adott közösség tagjai hogyan lépnek/léphetnek kapcsolatba egymással. Egy ilyen tervezési döntés lehet a moderálási folyamat mind inkább automatizmusokra történő felépítése is.³³

4. Az algoritmikus tartalom moderáció formái

Annak felderítése, hogy az egyes közösségimédia-platformok moderációs tevékenysége mögött pontosan milyen technikai megoldások húzódnak meg, közel sem triviális feladat. Ennek oka főként az, hogy kevés a nyilvánosan elérhető, megbízható információ a témában, hiszen az érintett vállalatok ezeket a legtöbb esetben üzleti titokként kezelik. Éppen ezért a helyzet feltáráshoz elengedhetetlen például a nyilvánosan elérhető jelentések, vagy éppen az olyan nem konvencionális források használata, mint az oknyomozó újságírás. Ezek felhasználása egyébként a vonatkozó szakirodalomban sem példa nélküli.³⁴

Az algoritmikus tartalommoderálás a közösségi média platformokon olyan technikák és eljárások összessége, amelyek célja a felhasználók által generált tartalom automatikus szűrése és felügyelete. Voltaképpen a már ismertetett *moderálási folyamat* au-

³¹ Marc Faddoul: COVID-19 is triggering a massive experiment in algorithmic content moderation. *Brookings*, 2020. április 28. <https://tinyurl.com/2p732haw>

³² James Grimmelman: The Virtues of Moderation. *Yale Journal of Law & Technology*, Vol. 17. (2015) 42–109. <https://doi.org/10.31228/osf.io/qwx5>

³³ Gergely Gosztonyi: Human and Technical Aspects of Content Regulation. *Erdélyi Jogélet*, Vol. 2., No. 4. (2022) 7–17.

³⁴ Robert Gorwa – Reuben Binns – Christian Katzenbach: Algorithmic content moderation: Technical and political challenges in the automation of platform governance. *Big Data & Society*, Vol. 7. No. 1. (2020) <https://doi.org/10.1177/2053951719897945>

tomatizálása ez, amelynek során a döntések adatbázisokkal való teljes, vagy részleges egyezései, esetleg gépi tanuló modellek által végzett osztályozás alapján történnek.

Napjainkban a legnagyobb közösségimédia platformok felhasználói bázisa bőven a milliárdos nagyságrendben van,³⁵ ezen felhasználók pedig ugyanekkora léptékben állítanak elő tartalmakat minden nap. Nem meglepő tehát, hogy a kizárólag moderátorok által végzett, vagy minden esetben általuk felülvizsgált ellenőrzési modell közelít a fenntarthatatlanhoz. A tartalommoderálás automatizálása ezzel szemben egy praktikusnak tűnő, jól skálázható megoldást jelent a közösségimédia-platformok számára.³⁶

Az alkalmazott megoldások rendkívül széles skálán mozognak mind összetettségben, mind pedig hatékonysági szempontok szerint. Közöttük talán a legfontosabb két kategória, a *hash* alapon működő, valamint a gépi tanulásra alapozott megoldások csoportja. Léteznek persze ennél történetileg régebbi, például kulcsszó alapon kereső eljárások is, de ezek képességei mai szemmel már annyira korlátozottak, hogy néhány speciális alkalmazási esetet leszámítva relevanciájukat is szinte teljes mértékben elvesztették.

2019-ben a Facebook nyilvánosan elérhetővé tette a PDQ, valamint a TMK + PDQF nevű, *hash* alapon működő megoldásait,³⁷ amelyek forráskódja mind a mai napig nyilvánosan elérhető.³⁸ A folyamat kulcsai a *hash* függvények. Ezek lényegében olyan algoritmusok, amelyek előállítják a fix hosszú reprezentációt olyan módon, hogy tetszőleges hosszúságú bitsort egy fix hosszúságúra képeznek le. Ilyen esetben kulcsfontosságú az, hogy amennyiben az eredeti bitsorozatban – például egy közösségimédia poszt szövegében – akár egyetlen karakter is megváltozik, akkor a belőle képzett, „hashelt” érték is módosuljon. Ezen felül azt is biztosítani kell, hogy nagyon alacsony legyen annak a valószínűsége, hogy két különböző szöveg hashelt értéke véletlenül megegyezzen.

Egy-egy tartalom *hash*ének kiszámítása után arról el kell döntenit, hogy az eredeti szöveg, kép, videó stb. megvalósít-e valamiféle szabály- vagy törvénysértést. Ilyen szabálysértés lehet bármi, ami a platform használatakor elfogadott ToS szerint ellentétes a közösség irányelveivel. Törvénysértő tartalmakra jó példa szerzői jogilag védett tartalmak ismételt feltöltése a felhasználó által sajátként. Ennek megállapításához léteznek olyan *hash*-adatbázisok, amelyek ismert problémás tartalmak *hash* értékeit tartalmazzák, és amelyeket az egyes szolgáltatók meg is oszthatnak egymással (*Shared Industry Hash Database*). Mivel a *hash* kiszámítása, és az ilyen jellegű adatbázisokban való keresés rendkívül gyors, ezért a módszer jól skálázódik még a közösségi médiára jellemző adatmennyiségek esetében is.

³⁵ Dave Chaffey: Global social media statistics research summary. *Smart Insights*, 2024. május 1. <https://tinyurl.com/sm74a5za>

³⁶ Joseph Seering – Tian Wang – Joon Youn – Geoff Kaufman: Moderator engagement and community development in the age of algorithms. *New Media & Society*, Vol. 21., No. 7. (2019) 1417–1443. <https://doi.org/10.1177/1461444818821316>

³⁷ Antigone Davis – Guy Rosen: Open-Sourcing Photo- and Video-Matching Technology to Make the Internet Safer. *Meta*, 2019. augusztus 1. <https://tinyurl.com/4p8e2d2f>

³⁸ Az informatikában, főleg a kriptográfiában a hashelés egy olyan eljárást jelent, amely során egy változó méretű bemenetből – például szövegekből vagy képekből – egy jellemzően rövidebb, fix hosszúságú reprezentációt állítanak elő.

A problémát itt a pontos egyezés kérdése jelentheti. Elvégre, például egy kép esetében elég csak egyetlen pixel színét megváltoztatni, és a kapott *hash* érték teljesen más lesz, mint amire az adatbázisban való keresés találatot jelezne. Egy pixel eltéréstől azonban egy eredetileg jogsértő tartalom jó eséllyel nem lesz kevésbé jogsértő; a jogszabályok ráadásul a platformot sok esetben arra is kötelezik, hogy az „azonos vagy egyenértékű” (lényegileg azonos) változatokat is eltávolítsa, amint arról az Európai Unió Bírósága döntött a Glawischnig-Piesczek kontra Facebook ügyben (C-18/18, ECLI:EU:C:2019:821).

Erre kínálnak megoldást az olyan *hash* függvények és adatbázisok, amelyek nem csak a pontos egyezéseket képesek előre jelezni, hanem az ún. lényegi egyezéseket is. A perceptuális, avagy „észlelési” *hash* függvények különbözősége abban áll, hogy két olyan tartalom esetében, amelyet az emberek hasonlóan észlelnek, ott ezek perceptuális *hash* értéke is azonos, vagy legalábbis nagyon hasonló lesz.³⁹ A módszer gyakorlatilag az emberi észleléssel nagyjából azonosnak mutakozó objektumokat (például képeket) hivatott ugyancsak azonosként, vagy *lényegileg azonosként* felismerni. Éppen ezért az ilyen algoritmusok alkalmasak arra, hogy felismerjék az olyan tartalmakat, amelyekre esetleg csak egy plusz vízjel került képek esetében, vagy detektálják, ha egy korábban gyűlöletkeltőnek talált szöveget valaki ismét feltölt néhány karakter megváltoztatásával. Valószínűsíthetően a Facebook fent említett megoldásai is ebbe a kategóriába tartoznak.⁴⁰

Egy másik fontos irányvonal a gépi tanuló algoritmusok, illetve a velük készített *modellek* alkalmazása. Ilyen esetben az ellenőrizendő tartalmak esetében nem korábbi, problémásként jelölt esetekkel való teljes vagy részleges egyezések felfedése a cél, hanem soha nem látott példák feletti döntések meghozatala (*prediktív* megközelítés). Ezt előre címkézett tanítóadatok segítségével lehetséges megvalósítani, amely a *felügyelt gépi tanulás*⁴¹ alapvető módszertana. Ilyenkor a gépi tanuló modell feladata az, hogy a tanítóadatokban látott mintázatok alapján olyan generalizálási képességre tegyen szert, amellyel tetszőleges új példákról jó hatékonysággal tudja eldönteni, hogy azok problémásak-e, vagy sem. Cél lehet ilyen esetben például az offenzív nyelvezet, a terrorizmust népszerűsítő, vagy éppen a pornográf tartalmak felismerése is. Utóbbira a Facebook is tett már dokumentált kísérletet.⁴²

Emellett léteznek ma már olyan kutatások is, amelyek a legmodernebb szabadon elérhető neurális hálózatokat alapul véve keresnek automatizált megoldást moderálási

³⁹ Hany Farid: An Overview of Perceptual Hashing. *Journal of Online Trust and Safety*. Vol. 1., No. 1. (2021) <https://doi.org/10.54501/jots.v1i1.24>; Moses Datar– Nicole Immerlica – Piotr Indyk et al.: Locality-sensitive hashing scheme based on p-stable distributions. In: *Proceedings of the Twentieth Annual Symposium on Computational Geometry*. New York, NY, ACM, 2004. 253–262. <https://doi.org/10.1145/997817.997857>

⁴⁰ Davis–Rosen (2019) i. m.

⁴¹ Vladimir Nasteski: An Overview of the Supervised Machine Learning Methods. *HORIZONS.B*, Vol. 4. (2017) 51–62. <https://www.doi.org/10.56726/IRJMETS51366>

⁴² Tarleton Gillespie: *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media*. New Haven, CT, Yale University Press, 2019. <https://doi.org/10.12987/9780300235029>

problémákra. Egy ilyen módszer lehet például a BERT modellek⁴³ betanítása a moderálással összefüggő feladatokra, például Reddit posztok kapcsán.⁴⁴ Ennek háttere az, hogy a Redditen számos közösség elkezdte a moderálási feladatokat és a feldolgozási logikát algoritmusokba beágyazni, ez pedig lehetőséget nyitott célzott tanítóadatok beszerzésére.⁴⁵ De ugyanígy ismerünk működő példákat a Wikipédia kapcsán is, ahol teljesen autonóm rendszerek ugyanígy működnek, mint a moderátorok támogatására szolgálók.⁴⁶ Szintén a kutatási szférából ismert olyan kísérlet is, amely ugyancsak pornográf tartalmak szűrésére többlépcsős, neurális hálózat alapú architektúra bevezetését javasolja.⁴⁷

5. A szólásszabadság és az információs ökoszisztéma

A modern közösségimédia-plattformok nem csupán technológiai vállalkozások, hanem az elmúlt évtizedben ezek váltak a társadalom alapvető kommunikációs infrastruktúrájának gerincévé is. Ezzel a szereppel azonban együtt jár az a képesség, hogy formálhatják és korlátozhatják a felhasználói interakciókat, diskurzusokat, és bizonyos esetekben a politikai részvételt vagy társadalmi szerepvállalást is.⁴⁸ Részben a leközölt tartalmak pusztá volumene, részben pedig törvényi kötelezettségeik miatt a már említetteknek megfelelően ma már nem támaszkodhatnak pusztán moderátorok tevékenységére a náluk megjelenő tartalmak szűrése során. Tevékenységükben az automatizált (algoritmikus) megoldások súlya éppen ezért folyamatosan növekszik.⁴⁹

⁴³ Jacob Devlin – Ming-Wei Chang – Kenton Lee – Kristina Toutanova: BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In: *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*. Minneapolis, Minnesota, Association for Computational Linguistics, 2019. 4171–4186. <https://doi.org/10.48550/arXiv.1810.04805>

⁴⁴ Qinglai He – Yili Hong – T. S. Raghu: Platform Governance with Algorithm-Based Content Moderation: An Empirical Study on Reddit. *Information Systems Research*, Vol. 36., No. 2. (2024) <https://doi.org/10.1287/isre.2021.0036>

⁴⁵ Eshwar Chandrasekharan – Savvas Mattia – Shagun Jhaver – Hannah Charvat – Amy Bruckman – Cliff Lampe – Jacob Eisenstein – Eric Gilbert: The internet’s hidden rules: an empirical study of Reddit norm violations at micro, meso, and macro scales. In: *Proceedings of the ACM on Human-Computer Interaction*, 2(CSCW) (2018) Article 32. <https://doi.org/10.1145/3274301>

⁴⁶ Aaron Halfaker – John Riedl: Bots and cyborgs: Wikipedia’s immune system. *Computer*, Vol. 45., No. 3. (2012) 79–82. <https://doi.org/10.1109/MC.2012.82>

⁴⁷ Dogus Karabulut – Cagri Ozcinar – Gholamreza Anbarjafari: Automatic content moderation on social media. *Multimedia Tools and Applications*, Vol. 82. (2023) 4439–4463. <https://doi.org/10.1007/s11042-022-11968-3>

⁴⁸ Sylvie Delacroix: Beware of “algorithmic regulation”. *SSRN Electronic Journal*, February 1, 2019. <https://doi.org/10.2139/ssrn.3327191>; Papp János Tamás: *A közösségi média szabályozása a demokratikus nyilvánosság védelmében*. Budapest, Wolters Kluwer, 2022.

⁴⁹ Delacroix i. m. 5–9.

5.1. Szólásszabadság

Az olyan platformok, mint például a Facebook vagy a Google tehát tagadhatatlanul jelentős hatalmat gyakorolnak a nyilvános kommunikáció felett. Ez a hatalom főként abban nyilvánul meg, hogy a platformok dönthetnek arról, milyen tartalmakat jelenítenek meg, milyen algoritmusokkal rendezik ezeket, és hogy sok esetben a platform mögött álló cégek döntenek el, mi minősül elfogadhatónak vagy elfogadhatatlannak. Ezzel azonban szükségszerűen befolyásolják azt is, hogy milyen információk érhetőek el a társadalom egésze számára, mely szereplők hallathatják a hangjukat a nyilvánosságban, illetve, hogy a szólásszabadság mikor és milyen mértékben korlátozódik.

Ez a fajta befolyás nem pusztán technikai vagy üzleti jellegű, hanem jelentős politikai következményei is vannak. A közösségimédia-platformok hatalma a kommunikáció felett azt jelenti, hogy ezek a cégek szabályozhatják a politikai diskurzust, elősegíthetik vagy akadályozhatják a politikai mozgalmakat, és akár közvetlenül befolyásolhatják a választási eredményeket is azáltal, hogy bizonyos információkat előnyben részesítenek – azaz több felhasználóhoz juttatnak el –, másokat pedig háttérbe szorítanak.⁵⁰

Fontos megjegyezni, hogy az ilyen platformok döntései kereskedelmi megfontolások alapján is történhetnek, amelyek gyakran nem esnek egybe a közérdekkel vagy a felhasználók érdekeivel. Ez felveti a kérdést, hogy szükséges lenne-e velük kapcsolatban szigorúbb szabályozásokat alkalmazni. Különösen fontos ez akkor, amikor az egy-egy ilyen szereplő által gyakorolt hatalom olyan mértékű, amelynek birtokában képesek alapvetően megváltoztatni a társadalmi kommunikáció dinamikáját. Ilyen típusú hatalommal a digitalizált világ elterjedése előtt *csakis államok, illetve egyházak rendelkeztek*. Jelenleg azonban a véleményformálásnak ezt a minden korábbinál hatékonyabb és átfogóbb képességét mindössze néhány nagy technológiai cég birtokolja, amelyek felhasználói bázisa országhatárokon túlnyúlva, lényegében az egész modern világra kiterjed.

A szólásszabadságra az algoritmusok esetleges tévedései is komoly hatással vannak. A gépi tanulási algoritmusok nagy mennyiségű adaton tanulnak, és céljuk, hogy felismerjék a mintázatokat, amelyek segítségével eldönthetik, hogy egy adott tartalom megfelel-e egy-egy platform szabályainak. Ilyenkor elkerülhetetlenül előfordulnak *fals pozitív*, és *fals negatív* esetek is. A szólásszabadságra gyakorolt hatás szempontjából a fals pozitív találatok különösen károsak lehetnek. Ilyen esetben az algoritmus tévesen távolít el egy egyébként ártalmatlan tartalmat, ezzel ok nélkül korlátozhatja a felhasználók véleménynyilvánítási jogát. Az ilyen típusú hibák különösen aggasztóak lehetnek a politikai vagy társadalmi diskurzusokban, ahol a szigorúbb moderálás a kritikus vélemények elhallgattatásához vezethet. Például, ha egy algoritmus valós ok nélkül töröl egy politikai tiltakozással kapcsolatos bejegyzést, azzal egyfajta cenzorként működik, és hozzájárul a véleménynyilvánítás korlátozásához.⁵¹

⁵⁰ Tomas Apodaca – Natasha Uzcátegui-Liggett: How Automated Content Moderation Works (Even When It Doesn't). *The Markup*, 2024. március 2. <https://tinyurl.com/f6f8j9v6>

⁵¹ Emma Llansó – Joris van Hoboken – Paddy Leerssen – Jaron Harambam: Artificial Intelligence, Content Moderation, and Freedom of Expression. *Transatlantic Working Group on Content Moderation Online*

Ezzel éppen ellentétesen, a fals negatív találatok egy eltérő kockázatot jelentenek. Amikor az algoritmus nem ismeri fel a szabálysértő tartalmat, azzal a káros információk, például dezinformáció, gyűlöletbeszéd vagy a szélsőséges, destruktív ideológiák szabadon történő terjedéséhez járul hozzá. Ez közvetlen veszélyt jelenthet a felhasználókra, és súlyosan torzíthatja a közvéleményt. Ráadásul az ilyen hibák alááshatják a közösségi média platformok hitelességét és a felhasználók beléjük vetett bizalmát is. A szólásszabadságra gyakorolt hatás persze a tartalomajánló rendszereken keresztül is érvényesül.

5.2. Véleménybuborékok és politikai polarizáció

A szólásszabadság formális keretei mellett azt is vizsgálni kell, hogyan torzulhat a nyilvános diskurzus az algoritmusok tartalom-rangsorolási logikája következtében.

A felhasználók elé kerülő tartalmak moderálása összetett folyamat, amely jóval túlmutat azon az egyszerű binárison, hogy egy tartalmat megtartanak vagy törölnek. A moderálás mára egy skálaszerű rendszerként működik, ahol a szankciók és beavatkozások változatos formában jelenhetnek meg. A döntés szólhat tartalomeltávolításról, de emellett számos olyan eszköz is létezik, amely láthatósági korlátozást alkalmaz. Ilyenek például a *shadow banning* (amikor a tartalom a közönség számára láthatatlan, de a közvétező számára nem), az algoritmikus háttérbeszorítás (*downranking*), a földrajzi alapú elérhetlenné tétel (*geoblocking*), vagy a személyes hírfolyamból való kiszorítás.

Emellett az is előfordulhat, hogy nem a tartalom, hanem a tartalomhoz kapcsolódó interakciók – például kommentelés vagy megosztás – kerülnek korlátozás alá, ami szintén befolyásolja annak elérhetőségét és társadalmi hatását. Ezek a moderációs gyakorlatok mind hozzájárulnak ahhoz, hogy milyen mértékben láthatók vagy észlelhetők egyes nézetek, és ez közvetlenül hat arra is, hogy a felhasználók milyen nézőpontokkal, politikai vagy kulturális álláspontokkal találkoznak online környezetükben.

Az algoritmikus tartalommoderálás és az ajánlási rendszerek együttes működése jelentős hatással lehet a társadalmi megosztottságra, különösen az úgynevezett visszhangkamrák (*echo chambers*) vagy szűrőbuborékok, esetleg véleménybuborékok (*filter bubbles*, *opinion bubbles*) kialakulásának elősegítésével.⁵² Ez a több néven is hivatkozott jelenség olyankor fordul elő, amikor a felhasználók főként olyan tartalmakkal találkoznak, amelyek megerősítik meglévő nézeteiket, miközben lényegében nem találkoznak ellentétes véleményekkel. Ilyen helyzet legegyszerűbben úgy jöhet létre, hogy például a közösségimédia-platformok tartalomajánló rendszerei hajlamosak kizárólagosan és személyre szabottan az adott felhasználó érdeklődésének, preferenciáinak megfelelő tartalmakat mutatni számára. Megjegyzendő, hogy a felhasználói preferenciákkal ellentétes nézetek nem teljes egészében válnak láthatatlanná. Gyakoriságuk azonban jelentősen csökken a hírfolyamokban és gyakran kedvezőtlen érzelmi

and Freedom of Expression, 2020. február 26. <https://tinyurl.com/4tcdtfr>

⁵² Eli Pariser: *The Filter Bubble: What the Internet Is Hiding from You*. New York, Penguin Press, 2011.; Papp János Tamás: Ajánlórendszerek és szűrőbuborékok. In: Koltay András (szerk.): *A vadnyugat vége? Tanulmányok az Európai Unió platformszabályozásáról*. Budapest, Wolters Kluwer, 2024. 231–259.

keretezésben jutnak el a közönséghez. Más szavakkal: a buborék nem hermetikusan zár, hanem torzítja a találkozás arányát és módját.⁵³

Ez a tartalom-rangsorolás aztán felerősíti egymást a megerősítési torzítással (*confirmation bias*⁵⁴) is. Ez arra a pszichológiai jelenségre utal, miszerint az emberek hajlamosak olyan információkat keresni, előnyben részesíteni és alapul venni, amelyek megerősítik a már elve meglévő vélekedéseiket, ellenben figyelmen kívül hagyni az ezzel ellentétes álláspontokat. Ez a kettő együttesen jelentősen befolyásolja a döntéshozatal és a véleményalkotást is. A téves bizonyosság ezenfelül közvetlenül hozzájárul a társadalmi megosztottság fokozódásához is.

A véleménybuborékok⁵⁵ létezése és a politikai polarizáció felerősödése szorosan összefügg, hiszen előbbiek közvetlenül elősegítik a politikai nézetek szélsőségesebbé válását és a társadalmi megosztottság növekedését.⁵⁶ A véleménybuborékok következtében a politikai diskurzus egyre inkább polarizálódik, mivel a különböző politikai csoportok egymástól elszigetelődnek, és kevésbé nyitottak az ellentétes nézetek meghallgatására vagy megértésére. Emellett a véleménybuborékokban a nézetek hajlamosak a szélsőségek felé tolni, ami egyúttal azt is jelenti, hogy a kompromisszum keresése egyre inkább kiveszik, helyét pedig a egyre radikálisabb nézetek propagálása veszi át.

E mögött persze az is fellelhető, hogy a közösségi médiában a megosztó (akár szélsőséges) tartalmak jellemzően több interakciót generálnak, mint a konzolidáltak. Az interakciók nagyobb figyelmet is eredményeznek, ami miatt pedig a figyelmi idő maximalizálásra törekvő vállalatoknak paradox módon érdekük, hogy az ilyen tartalmakat népszerűsítsék. Ennek háttérében az áll, hogy a figyelemalapú gazdaságban⁵⁷ a szolgáltatók célja, hogy a felhasználóknak a platformjukon töltött idejét maximalizálják.

6. Az automatizált moderáció közvetlen hatásai a felhasználókra

6.1. Dezinformáció terjedése

A buborékhatás következményeként jelentkezhet a dezinformációk erősödése is, különösen, ha az algoritmusok olyan tartalmakat részesítenek előnyben, amelyek fokozzák

⁵³ Eli Pariser: Did Facebook's Big Study Kill My Filter Bubble Thesis? *Wired*, 2015. május 7. <https://www.wired.com/2015/05/did-facebooks-big-study-kill-my-filter-bubble-thesis/>

⁵⁴ Joshua Klayman: Varieties of Confirmation Bias. *Psychology of Learning and Motivation*, Vol. 32. (1995) 385–418. [https://doi.org/10.1016/S0079-7421\(08\)60315-1](https://doi.org/10.1016/S0079-7421(08)60315-1)

⁵⁵ A. K. Saxena: Beyond the Filter Bubble: A Critical Examination of Search Personalization and Information Ecosystems. *International Journal of Intelligent Automation and Computing*, Vol. 2., No. 1. (2019) 52–63.

⁵⁶ Emily Kubin – Christian von Sikorski: The Role of (Social) Media in Political Polarization: A Systematic Review. *Annals of the International Communication Association*, Vol. 45., No. 3. (2021) 188–206. <https://doi.org/10.1080/23808985.2021.1976070>; David Hartmann – Sonja Wang – Lena Pohlmann – Bettina Berendt: A systematic review of echo chamber research: comparative analysis of conceptualizations, operationalizations, and varying outcomes. *Journal of Computational Social Science*, Vol. 8., Article n. 52. (2025) <https://doi.org/10.1007/s42001-025-00381-z>

⁵⁷ Thomas H. Davenport – John C. Beck: The Attention Economy. *Ubiquity*, May 2001. <https://doi.org/10.1145/376625.376626>

a figyelmet vagy felháborodást. A dezinformáció egyik kitüntetett esete a hamis hírek (*fake news*), valamint a mostanában feltörekvően lévő – a generatív mesterséges intelligenciára is nagyban támaszkodó – *deepfake*-ek terjedése. Az *álhírek* olyan tudatosan terjesztett félrevezető információk, amelyek megjelenhetnek gazdasági haszonszerzés, kattintásvadászat, szórakoztató vagy éppen szándékolt károkozás (*hoax*) motivációjával is, miközben a politikai vagy geopolitikai manipulációt célzó dezinformáció egy különösen veszélyes formájuknak számít.⁵⁸ A politikai dezinformáció szándékos terjesztése a pszichológiai hadviselés egyik eszköze is lehet. A mindennapi beszédben álhírként emlegetünk minden olyan hamis információt, amely széles körben eljut a nyilvánossághoz. A *deepfake* egy olyan technológia, amely mesterséges intelligencia, konkrétan gépi tanulás segítségével képes hamisított képeket, videókat vagy hangfelvételeket készíteni, amelyek azonban valóságúnak tűnnek.⁵⁹

A tartalommoderálás, és az erre alkalmazott algoritmusok egyik fő célja éppen az ilyen és ehhez hasonló dezinformáció terjedésének megakadályozása. Gyakran előfordul azonban, hogy ezek az automatizált rendszerek nem képesek időben felismerni vagy eltávolítani a hamis(itott) információkat. Ez különösen problematikus a gyorsan terjedő tartalmak esetében (*viral content*), mivel az algoritmusok, amelyek az elköteleződés (lájkok, megosztások, kommentek) alapján rangsorolják a tartalmakat, hajlamosak az ilyen típusú információkat gyorsan és széles körben terjeszteni. Ezért szélsőséges esetben még azelőtt felhasználók tömegeihez juthat el egy-egy problémás hír, vagy videó, hogy a vele kapcsolatos moderálási döntés megszületne. Ennek hátterében az áll, hogy noha a moderálás maga nagyrészt ma már automatizált, még mindig nagyban támaszkodik a felhasználóktól érkező bejelentésekre is, amelyekhez azonban sokszor idő kell. Automatikus törlés vagy korlátozás hiányában pedig a dezinformáció sokkal gyorsabban elérheti a közönséget, mint ahogy a moderátorok beavatkozhatnának a folyamatba.

A mesterséges intelligencia rendszerek szabályozását rendező uniós rendelet (*AI Act*) kimondja,⁶⁰ hogy a szintetikusan létrehozott vagy „*deepfake*” módon jelentősen módosított kép-, hang-, videó- és szövegtartalmakat egyértelmű, jól látható jelöléssel kell ellátni, hogy a felhasználók tisztában legyenek a látott, hallott vagy olvasott információ valóságtartalmával. Előírja továbbá, hogy a jelölés „géppel olvasható”, azaz automatikusan detektálható módon is megtörténjen. Ezt a gyakorlatban érvényesíteni viszont közel sem triviális feladat. Különösen, ha figyelembe vesszük, hogy a dezinformáció terjesztésében érdekelteknek magától értetődően nem érdeke az ilyesfajta együttműködés, inkább ennek megkerülése.

⁵⁸ Howard Tumber – Silvio Waisbord (szerk.): *The Routledge Companion to Media Disinformation and Populism. 1. edition.* London, Routledge, 2021.

⁵⁹ Anupama Chadha – Vaibhav Kumar – Sonu Kashyap – Mayank Gupta: Deepfake: An Overview. In: *Proceedings of Second International Conference on Computing, Communications, and Cyber-Security.* Springer, Singapore, 2021. 557–566. https://doi.org/10.1007/978-981-16-0733-2_39; Aczél Petra – Veszelszki Ágnes (szerk.): *Deepfake: a valótlán valóság.* Budapest, Gondolat, 2023.; Gosztonyi Gergely – Lendvai Gergely: Deepfake és dezinformáció – Mit tehet a jog a mélyhamisítással készített álhírek ellen? *Médiakutató*, 2024/1. 41–49. <https://doi.org/10.55395/MK.2024.1.3>

⁶⁰ Regulation (EU) 2024/1689 of the European Parliament and of the Council laying down harmonised rules on Artificial Intelligence (*AI Act*), HL L 295, 2024.09.12., Art. 50 (2)–(4).

További probléma, hogy az algoritmusok gyakran nem képesek megfelelően értelmezni a tartalom kontextusát, ami azt eredményezheti, hogy a dezinformációval szembeni küzdelem során valós és releváns információk is eltávolításra kerülnek.⁶¹

A kontextusértelmezés persze nemcsak a gépi, hanem az emberi moderálásnak is tartós Achilles-sarka. A gyűlöletbeszéd-szabályozásban például éppen ezért született meg az ENSZ által támogatott Rabat-akcióterv (*Rabat Plan of Action*) hatpontos tesztje, amely a szándék, a tartalom, a terjesztés közege, a beszélő helyzete, a célcsoport kiszolgáltatottsága és a beszéd valószínű hatása alapján mérlegeli, mikor indokolt a korlátozás. A gyakorlat azonban azt mutatja, hogy még e keretrendszer alkalmazása mellett is gyakran előfordulnak téves eltávolítások, mert a kontextus (pl. idézés, művészi cél vagy ellenbeszéd) nem mindig dönthető el a rendelkezésre álló információkból. Ez jól illusztrálja: ha az emberi szakértők is vitatkoznak egy tartalom besorolásán, akkor a tisztán algoritmikus döntéshozás szükségszerűen még nagyobb hibaarányal működik.

Mindez különösen igaz az olyan helyzetekben, amikor a hamis információk és a valós tények között csak árnyalatnyi különbségek vannak, vagy ahol szarkazmus, ironia, vagy kulturálisan meghatározott speciális kifejezések vagy értelmezési lehetőségek jelennek meg.

6.2. Egyenlő bánásmód és algoritmikus torzítás

A technológia fejlődése során nyilvánvalóvá vált, hogy az automatizált rendszerek gyakran diszkriminatívan járnak el, különösen a marginalizált közösségekkel szemben. Ezek a torzítások sok esetben az algoritmusok fejlesztéséhez használt tanítóadatokban gyökereznek, amelyek implicit módon magukban hordozhatják a társadalomban már meglévő előítéleteket és sztereotípiákat.

Bármilyen gépi tanulásra alapozott megoldás esetében igaz, hogy az algoritmus hatékonysága nagymértékben függ a tanítóadatoktól. Ha ezek az adatok torzítanak (*biased*) vagy hiányosak, az algoritmusok is ezekkel a hibákkal fognak működni. Az adatokban jelenlévő torzítások lehetnek explicitek és implicitek is, mely utóbbiak korrigálása sokkal körülményesebb, ha megoldható egyáltalán.⁶² Például, ha egy moderálási algoritmus betanítása során aránytalanul sok esetben kerülnek eltávolításra egy adott kisebbséghez tartozó felhasználók bejegyzései, akkor az algoritmus később is nagyobb valószínűséggel fogja eltávolítani az ilyen bejegyzéseket, még akkor is, ha azok nem sértik a közösségi irányelveket. Ezzel szemben más közösségek hasonló tartalmai érintetlenek maradnak, ami pedig közvetlenül egyfajta egyenlőtlen bánásmódhoz vezet. A közel-múlt empirikus és áttekintő kutatásai szisztematikusan rámutatnak arra, hogy a gépi tanuláson alapuló rendszerek a fejlesztés minden szakaszában (az adatok gyűjtésétől a modellek finomhangolásáig) hordozhatnak és felnagyíthatnak társadalmi előítéleteket. 2022-ben az Európai Unió Alapjogi Ügynöksége (FRA) részletesen dokumentálta, hogy

⁶¹ Delacroix (2019) i. m. 3–6.

⁶² Ferry Hoitsma – Guillermo Nápoles – Çiğdem Güven et al.: Mitigating implicit and explicit bias in structured data without sacrificing accuracy in pattern classification. *AI & Society*, Vol. 40. (2024) <https://doi.org/10.1007/s00146-024-02003-0>

a pontatlan vagy torz tanítóadatok „jelentős és tartós diszkriminációt” idézhetnek elő, ha a szolgáltatók nem végeznek előzetes célzott ellenőrzéseket és utólagos validálást is.⁶³ Több, az egészségügyben alkalmazott gépi tanulási modellekre fókuszáló összegzés is hasonló következtetésre jutott.⁶⁴ Az egészségügyi rendszereket vizsgáló metaanalízis szerint, amely 30 tanulmányt összegez, a klinikai gépi tanulási modellek több mint felében kedvezőtlenebb predikciók várhatók fekete vagy alacsony jövedelmű páciensekre, amelyek az algoritmikus torzításon kívül máshogyan nem indokolhatók.⁶⁵

A fentiekre válaszul a fejlesztők egyre gyakrabban alkalmaznak olyan adat- vagy modell-szintű technikákat (például előzetes újrasúlyozás, „fair” veszteségfüggvény, utólagos küszöbhangolás), amelyek célja a csoportok közötti statisztikai különbségek mérséklése (*fairness intervention*). E módszerek azonban kompromisszumot követelnek pontosság, stabilitás és jogi megfelelés között. Ez lényegében azt jelenti, hogy a rendelkezésre álló tanítóadatok mesterséges manipulálása a torzítások kiküszöbölése érdekében csak akkor lehet valóban hatékony, ha a hatások elemzése maga is többlépcsős, külön fejlesztési folyamat. Ennek hiányában a torzítások kiküszöbölése nem garantált: lényegében csak a modellek általános teljesítménye csökken, de a torzítások nem szűnnek meg. Ezt jól illusztrálja egy nemrégiben megjelent szisztematikus áttekintés, amely 50 pénzügyi, HR- és igazságügyi rendszert elemzett. Az adatokból látható, hogy az elnagyolt, nem kellőképpen auditált beavatkozások a modellek tanítóadataiba az esetek harmadában csak tovább rontották a kisebbségi csoportokra vonatkozó modell-predikciókat, negyedüknél egyáltalán nem javították azt, és csak kevesebb mint a felüknél érték el valódi torzításmentes állapotot.⁶⁶ A probléma érdemi kezeléséhez a technológiai cégeknek rendszeres, független adatauditokat kellene bevezetniük, amelyek nemcsak a tanítóadatok sokszínűségét és pontosságát vizsgálják, hanem a főbb metszetekre kiterjedő érzékenységi teszteléseket is elvégzik, azaz a különböző változók (bőrszín, nem stb.) együttes hatását is vizsgálják. Emellett fontos, hogy a tartalommoderálás során az emberi moderátorok jobban kiegészítsék és korrigálják az algoritmusok által hozott döntéseket (pl. *Reinforcement Learning with*

⁶³ Christiane Wendehorst: *Bias in Algorithms: Artificial Intelligence and Discrimination*. Luxembourg, Publications Office of the European Union, 2022. <https://tinyurl.com/5t5evc6h>

⁶⁴ Michael Colacci – Yu Qing Huang – Gemma Postill – Pavel Zhelnov – Orna Fennelly – Amol Verma – Sharon Straus – Andrea C. Tricco: Sociodemographic bias in clinical machine learning models: a scoping review of algorithmic bias instances and mechanisms. *Journal of Clinical Epidemiology*, Vol. 178., 111606. (2025) <https://doi.org/10.1016/j.jclinepi.2024.111606>

⁶⁵ Syed Ali Haider – Sahar Borna – Cesar A. Gomez-Cabello – Sophia M. Pressman – Clifton R. Haider – Antonio Jorge Forte: The Algorithmic Divide: A Systematic Review on AI-Driven Racial Disparities in Healthcare. *Journal of Racial and Ethnic Health Disparities*, 2024. 12. 18., online ahead of print. <https://doi.org/10.1007/s40615-024-02237-0>

⁶⁶ Feng Chen – Liqin Wang – Julie Hong – Jiaqi Jiang – Li Zhou: Unmasking bias in artificial intelligence: a systematic review of bias detection and mitigation strategies in electronic health record-based models. *Journal of the American Medical Informatics Association*, Vol. 31. No. 5. (2024), 1172–1183. <https://doi.org/10.1093/jamia/ocae060>

Human Feedback – RLHF⁶⁷). Ez különösen fontos olyan esetekben, ahol a kontextus meghatározó egy-egy döntés jogossága kapcsán. Emellett persze a nagyobb átláthatóság és elszámoltathatóság is elengedhetetlen lenne a felhasználói bizalom megszlárdítása érdekében. A legtöbb közösségimédia-platform ma üzleti titokként kezeli a moderálásra alkalmazott algoritmusokat, amely gyakorlat azonban – még ha üzletileg védhető is – jelentősen hátráltatja a velük szembeni bizalom kiépítését, és tekintettel az ilyen cégek szólásszabadságra gyakorolt hatására, több szempontból aggályos gyakorlat is.⁶⁸

6.3. Mentális és pszichológiai hatások

Az algoritmikus tartalommoderálás ténye és megnyilvánulási formái komoly negatív pszichológiai hatást fejthet ki a felhasználókra, különösen akkor, ha azok gyakran szembesülnek indokolatlan hibaüzenetekkel (például egy bejelentés során), vagy nem megfelelően kommunikált moderációs döntésekkel.

Amikor a felhasználók azt tapasztalják, hogy tartalmaikat eltávolítják vagy blokkolják anélkül, hogy erre világos magyarázatot kapnának, az értelemszerűen frusztráció kialakulásához fog vezetni (vö. fals pozitív esetek). A felhasználók számára átláthatatlan és kiszámíthatatlan lehet a fellebbezési folyamat is, amely ráadásul gyakran időigényes és bonyolult, és nem is mindig vezet kielégítő eredményre.⁶⁹

A fentiekkel kapcsolatban érdemes persze megemlíteni, hogy a DSA 17. cikke immár kifejezetten kötelezi a platformokat, hogy minden eltávolításról vagy korlátozásról szóló döntést indokolással küldjenek a felhasználónak, és biztosítsanak könnyen elérhető, pártatlan fellebbezési csatornát is. Ennek célja éppen az, hogy csökkentse a nehezen átlátható moderációs folyamatok, döntések okozta frusztrációt.

A szorongás szintén jelentős pszichológiai következmény lehet, különösen azok számára, akik rendszeresen használják az ilyen platformokat kommunikációra, önkifejezésre vagy akár üzleti célokra. A tartalom moderálásával kapcsolatos állandó bizonytalanság, valamint az attól való félelem, hogy egy fontos bejegyzést vagy fiókot eltávolítanak, komoly stresszfaktorként is jelentkezhet. A tartalmak rendszeres eltávolítása vagy a fiókok blokkolása súlyosan befolyásolhatja a felhasználók mentális egészségét, különösen azoknál, akik online közösségek részei, vagy akiknek online jelenléte alapvető fontosságú a munkájukhoz vagy társadalmi kapcsolataik kiépítéséhez, esetleg fenntartásához.

Mindezek fényében kulcsfontosságú, hogy a platformok a következő években ne pusztán jogi megfelelési feladatként tekintsék a DSA-kötelezettségekre, hanem – a felhasználók mentális jóllétét is szem előtt tartva – transzparens, gyors és empatikus

⁶⁷ Guangli Li – Randy Gomez – Keisuke Nakamura – Bo He: Human-Centered Reinforcement Learning: A Survey. *IEEE Transactions on Human-Machine Systems*, Vol. 49., No. 4. (2019) 337–349. <https://doi.org/10.1109/THMS.2019.2912447>

⁶⁸ Zödi Zsolt: *Platformjog*. Budapest, Ludovika, 2023.

⁶⁹ Kris Holt: Meta's Oversight Board Raises Concerns Over Automated Moderation of Hate Speech. *Engadget*, January 23, 2024. <https://tinyurl.com/b39t8uuk>

kommunikációs protollokat alakítsanak ki. E protolloknak egyaránt tartalmazniuk kell(ene) a döntések laikusok számára is érthető indoklását, a hibás szankciók gyors visszafordítását, valamint az érintettek pszichológiai támogatására szolgáló tájékoztató anyagokat. Ha a moderációs lánc minden szintjén megjelenik ez a szemlélet, az nemcsak a felhasználói bizalom helyreállítását segítheti, hanem hosszú távon az online közösségi terek egészségesebb, kevésbé polarizált működéséhez is hozzájárulhat.

7. Konklúzió

A mesterséges intelligencia alapú tartalomszűrő algoritmusok elterjedése jelentős mérföldkönek számított a közösségimédia-platformok működésében, ám alkalmazásuk számos kihívással jár. Az ilyen rendszerek számos előnyös tulajdonsággal bírnak, hiszen képesek hatalmas mennyiségű tartalmat kezelni rövid idő alatt, ami a platformok növekvő felhasználói bázisát tekintve szükséges is. Ezek az algoritmusok azonban gyakran nem képesek megfelelően kezelni a tartalom kontextusát, ami számos téves döntéshez vezethet, és ez komolyan veszélyezteti a felhasználók jogait, különösen a szólásszabadságot.

Az algoritmusok működése szorosan összefügg a felhasznált adatok minőségével és sokszínűségével. Amennyiben a gépi tanulási modellek nem megfelelően kiegyensúlyozott adatokon alapulnak, az eredmények diszkriminatívak lehetnek bizonyos társadalmi csoportokkal szemben. Ez különösen problematikus a marginalizált közösségeket illetően, akik emiatt gyakran szembesülhetnek az algoritmusok okozta torzításokkal. A technológiai cégek felelőssége, hogy olyan adatokat és algoritmusokat alkalmazzanak, amelyek biztosítják a tartalom moderálásában az egyenlő bánásmód érvényesülését.

További aggodalomra ad okot a véleménybuborékok kialakulása, amelyek az algoritmusok személyre szabott tartalomajánló rendszereinek melléktermékei. Ezek a rendszerek megerősítik a felhasználók meglévő nézeteit, miközben kizárják az alternatív álláspontokat. A folyamat elősegíti a politikai és társadalmi polarizációt, mivel a felhasználók egyre inkább csak olyan tartalmakkal találkoznak, amelyek megerősítik saját világnézetüket. A szűkített perspektíva veszélye abban rejlik, hogy a társadalmi párbeszéd sokszínűsége elvész, ami súlyos következményekkel járhat a demokratikus folyamatokra nézve.

Ahhoz, hogy ezek a kihívások kezelhetők legyenek, elengedhetetlen a tartalommoderálás szabályozásának korszerűsítése, az ilyen rendszerek transzparenciája, és az emberi tényező hangsúlyosabbá tétele a folyamatban. Az algoritmusok mellett a moderátorok szerepe sem veszíthet jelentőségéből, hiszen az emberi ítélőképeség és az empátia továbbra is nélkülözhetetlen a szürke zónák kezelése során. A szabályozók, a felhasználói civil/társadalmi szereplők és a technológiai cégek közötti együttműködés javítása is kulcsfontosságú, hogy az algoritmusokkal való moderálás fenntartható és igazságos legyen, miközben megőrzi az online terek sokszínűségét és az általános szólásszabadságot.

VARIA

– BESZÁMOLÓ –

VI. FOGYASZTÓVÉDELMI AKTUALITÁSOK
– A DIGITALIZÁCIÓ, A MESTERSÉGES INTELLIGENCIA
FOGYASZTÓVÉDELMI KIHÍVÁSAI

A Budapesti Gazdaságtudományi Egyetemen 2022-ben alakult, több egyetem és tudományág képviselőit tömörítő Fogyasztóvédelmi Kutatócsoport¹ 2025. május 28-án tartotta VI. *Fogyasztóvédelmi aktualitások – a digitalizáció, a mesterséges intelligencia fogyasztóvédelmi kihívásai* című konferenciáját, amelynek keretében a témába vágó előadások mellett sor került a huszonkét szerző – Barna Sarolta, Bencsik András, Fazekas Judit, Fekete-Frojimovics Zsófia, Filó Mihály, Frivaldszky János, Hamar Farkas, Hámori Antal, Hernádi Júlia, Kasza Gyula, Koller Edit Mária, Koltay András, Lapsánszky András, Nagy Dániel, Nemes Imre, Rigó Csaba Balázs, Sándor István, Szabó Edit, Szőke Andrea, Totth Gedeon, Varga József, Veres Zoltán – által megalkotott tizenhárom fejezetet és Vékás Lajos akadémikus Bevezetését tartalmazó, *Fogyasztóvédelem – fogyasztóvédelmi jog* című könyv bemutatására is.² A megjelenteket Kiss Kornélia dékán köszöntötte, aki kiemelte, hogy az oktatás, kutatás és társadalmi felelősségvállalás kéz a kézben kell, hogy járjanak. Ezért a fogyasztóvédelem nemhogy nem akadály, hanem lehetőség, amely segíthet abban is, hogy

¹ A Kutatócsoporttal kapcsolatos további információk a <https://uni-bge.hu/hu/kvik/fogyasztovedelmi-kutato-csoport> oldalon érhetők el. A korábbi konferenciákkal kapcsolatban ld. pl. Veres Zoltán: *Fogyasztóvédelmi aktualitások – interdiszciplináris megközelítésben. Iustum Aequum Salutare*, 2023/1. 203–206.; Veres Zoltán: *Beszámoló a Budapesti Gazdasági Egyetemen működő Fogyasztóvédelmi Kutatócsoport „II. Fogyasztóvédelmi aktualitások” című konferenciájáról. Jogelméleti Szemle*, 2023/2. 121–124.; Hámori Antal – Veres Zoltán: *Fókuszban a fogyasztókkal szembeni tisztességtelen kereskedelmi gyakorlatok tilalma – Beszámoló a BGE Fogyasztóvédelmi Kutatócsoport III. tudományos konferenciájáról. Jog – Állam – Politika*, 2024/1. 211–214.; Veres Zoltán – Hámori Antal: *IV. Fogyasztóvédelmi aktualitások című konferencia a Budapesti Gazdasági Egyetemen. De iurisprudentia et iure publico*, 2025/1–2. 103–106.

² Hámori Antal – Hamar Farkas – Neszmélyi György Iván – Veres Zoltán (szerk.): *Fogyasztóvédelem – fogyasztóvédelmi jog*. Budapesti Gazdaságtudományi Egyetem, Budapest, 2025. <https://doi.org/10.29180/978-615-6886-13-2> (a teljes könyv lektora: Vékás Lajos; a fejezetek lektorai: Árva Zsuzsanna, Balázs István, Barta Judit, Benke József, Fézer Tamás, Gellén Klára, Gombos Katalin, Győri Zoltán, Józsa László, Karsai Krisztina, Lentner Csaba, Nagy Zoltán, Rekettye Gábor).



fenntarthatóan tudjuk működtetni rendszereinket. A könyvről szólva elmondta: az nem csak egy tudományos mű, hanem közös gondolkodás eredménye, hiszen több egyetem és hatóság szakemberei írták.

Ezt követően Hámori Antal felelős szerkesztő, a Kutatócsoport vezetője arra mutatót rá, hogy a könyv egy gazdag körutazásra invitálja az olvasót, és kifejezte reményét, hogy az az elmélet és gyakorlat számára is hasznosítható, hiszen megírása során nagy hangsúlyt helyeztek a joggyakorlat figyelemmel kísérésére és a jogalkalmazó szervezetekkel való együttműködésre. Koller Edit ny. főosztályvezető (PVKH) saját tapasztalatit megosztva elmondta, hogy a könyv fejezetének írása számára egyfajta tanulási folyamat is volt, lehetőséget adva olyan összefüggések felismerésére, amelyeket a napi rutin során nem vett észre. Lapsánszky András elnökhelyettes (NMHH) az általános és speciális szakterületek együttműködésének fontosságára mutatott rá, mivel a fogyasztóvédelemben fontos az új eszközök meghonosítása. E gondolatmenet mintegy folytatásaként Hernádi Júlia irodavezető-helyettes (GVH) a digitalizáció kihívásaival kapcsolatban bemutatta, hogy a GVH számos eljárást indított nagy technológiai változatok ellen. Miklós-Illés Gergely főosztályvezető (PVKH) hozzászólásában kiemelte: erős elméleti tudás nélkül szinte lehetetlen megfelelő jogalkalmazást folytatni; a digitalizáció jelentette kihívások fontosságát alátámasztja, hogy saját területén is az online térhez kötődtek legnagyobb számban a hatósághoz érkezett beadványok.

A könyvbemutatót a konferencia-előadások követték. Sándor István tanszékvezető, egyetemi tanár (ELTE ÁJK Római Jogi és Összehasonlító Jogtörténeti Tanszék) *A digitalizáció kihívásai a fogyasztók magánjogi védelme területén* címet viselő előadásában kiindulópontként rögzítette: napjaink adatvezérelt gazdaságában még fokozottabban érvényesül az információs aszimmetria. Ami a fogyasztói szerződéseket illeti, a digitális térben nincs valódi alkupozició, tárgyalási lehetőség, de a megvásárolni kívánt termék, szolgáltatás alapos megvizsgálása is akadályokba ütközik. Sokszor azt sem könnyű kideríteni, valójában kivel kötünk szerződést. E tényezők okán a fogyasztók kiszolgáltatottsága tovább nő, ami felveti azt a kérdést, hogy a klasszikus polgári jogi eszközök mennyiben biztosítanak hatékony védelmet részükre. Ezt felismerve a szabályozás (DSA és DMA) bevezette a platformfelelősség intézményét. Mindazonáltal a vázolt tényezők magával a jogalkotással kapcsolatban is számos dilemmát felvetnek, így például azt, hogy egyáltalán ember fogja-e alkotni a szabályokat? De adódik a klasszikus kérdés is: paternalista módon vagy a szerződési szabadság talaján állva alkossuk meg a megfelelő szabályozást?

Fazekas Judit prodékán, egyetemi tanár (SZE DF ÁJK Polgári Jogi és Polgári Eljárásjogi Tanszék): *Ki fizeti a révést? – avagy a mesterséges intelligencia által okozott károkért való magánjogi felelősség viselése* című előadásában először az uniós jogi keretéről szólva ismertette az AI Act³ főbb prioritásait, majd vázolta a felelősségi

³ Az Európai Parlament és a Tanács (EU) 2024/1689 rendelete (2024. június 13.) a mesterséges intelligenciára vonatkozó harmonizált szabályok megállapításáról, valamint a 300/2008/EK, a 167/2013/EU, a 168/2013/EU, az (EU) 2018/858, az (EU) 2018/1139 és az (EU) 2019/2144 rendelet, továbbá a 2014/90/EU, az (EU) 2016/797 és az (EU) 2020/1828 irányelv módosításáról. HL L, 2024/1689, 2024.7.12. <https://eur-lex.europa.eu/legal-content/HU/TXT/?uri=CELEX:32024R1689>

ikerirányelvek rendszerét. Az új termékfelelősségi irányelv kapcsán kiemelte, hogy a termék fogalmába tartoznak a szoftver és a digitális fájlok is. Áttérve a magyar felelősségi szabályokra, az előadó a nagy kockázatú AI rendszerekre a veszélyes üzemi felelősség szabályait tekinti alkalmazandónak, hozzátéve, hogy a termékfelelősségi irányelv implementációja törvényt módosítást igényel. A szabályozás alakítása során nem kerülhető meg az AI területén való jártasság követelményének átgondolása sem.

Frivaldszky János tanszékvezető, egyetemi tanár (PPKE JÁK Jogbölcseleti Tanszék) *A neurotechnológia kétes szerepe a fogyasztói kereskedelemben* címmel tartotta meg referátumát. Ebben felhívta a figyelmet, hogy a fogyasztóvédelem klasszikus emberképe (jól informált, autonóm szerződő fél) már nem létezik, ugyanis a transzhumanizmus – eliminálva a szellemi és pszichológiai dimenziót – neuronok, hormonok (mint a dopamin⁴) szintjére süllyesztette az embert. Így a fogyasztás könnyen vallásos áhitattá válhat, amit a neuromarketing igyekszik is felerősíteni, kihasználni.

Bencsik András egyetemi docens (ELTE ÁJK Közigazgatási Jogi Tanszék) *Korszakunk digitalizációs kihívásainak fogyasztóvédelmi vonatkozásairól* szólva felhívta a figyelmet az informatikai infrastruktúra fontosságára, ennek hiányában – mint számos példa mutatja (ld. elektronikus anyakönyv bevezetése) – a digitalizáció nem kivitelezhető. Természetesen vannak olyan tevékenységek, amelyek könnyebben, mások nehezebben digitalizálhatók. Bár a digitalizáció látszólag decentralizációt eredményez (hiszen az ügyek bárholnán intézhetők), valójában centralizáló, hiszen az adatok központosítását feltételezi. Az előadó szolt a mesterséges intelligencia által felvetett szerzői jogi kihívásokról is (például a tanuló adatok felhasználása, vagy a szerzői mű minősítésének kérdései), majd rámutatott arra, hogy a mesterséges intelligencia kapcsán gyakran emlegetett „black box effektus” az emberi jogalkalmazás során is jelen van.

Tóth András tanszékvezető, egyetemi docens (KRE ÁJK Digitális és Technológia Jogi Tanszék), elnökhelyettes (GVH), elnök (Versenytanács): *A digitális fogyasztóvédelem széttöredezett európai szabályozása* című előadásában szemléletesen bemutatta az egyes területeken fennálló szabályozási párhuzamokat, amelyek egy rendkívül bonyolult jogi keretrendszer eredményeznek. A probléma megoldását a Bizottság a szubszidiaritás ellenében képzeli el. Ezzel kapcsolatban az előadó fenntartásait hangoztatta, hiszen a központosítás egyebek mellett a nemzeti hozadék elvesztését is eredményezné, pedig a magyar versenyhatóság számos eredményt ért el a digitális fogyasztóvédelem területén, köszönhetően az elmúlt tíz esztendőben folytatott eljárások során szerzett tapasztalatoknak is.⁵

⁴ A boldogságérzetet alapvetően négy hormonnak, a dopaminnak, a szerotoninnak, az oxitocinnak és az endorfinoknak köszönhetjük. Amikor ezek felszabadulnak a vérben, hirtelen boldognak érezzük magunkat. Perei Dóra: Úgy tűnik alábecsültük a dopamin és a szerotonin szerepét életünk alakításában. *Rakéta*, 2021. 01. 01. <https://tinyurl.com/4pxdjc9t> Kishida szerint a dopamin és a szerotonin kulcsszerepet játszik a döntéshozatalban, valahányszor elbizonytalanodunk, illetve cselekvésre is ösztönöz. Ld. Dan Bang – Kenneth T. Kishida – Terry Lohrenz – Jason P. White – Adrian W. Laxton – Stephen B. Tatter – Stephen M. Fleming – P. Read Montague: Sub-second Dopamine and Serotonin Signaling. *Human Striatum during Perceptual Decision-Making. Neuron*, Vol. 108., No. 5. (2020) <https://doi.org/10.1016/j.neuron.2020.09.015>

⁵ Ld. pl. Google, Alza.hu, Facebook, TikTok, Booking.

Veres Zoltán vármegyei jegyző (KEVÖ), megbízott oktató (PPKE JÁK Pénzügyi Jogi Tanszék) *Egyes fogyasztóvédelmi kérdések a mesterséges intelligencia pénzügyi szektorban történő alkalmazása kapcsán* című előadásában bevezetésként a mesterséges intelligencia fogalmának néhány megközelítési lehetőségét ismertette, majd vizsgálta annak alkalmazási lehetőségeit, előfordulásait először a mindennapokban, majd a pénzügyi szektorban. Megállapította, hogy a hazai pénzügyi szektor szereplői alkalmaznak ugyan MI-rendszereket, de egyelőre inkább a kísérletezés és útkeresés, további felhasználási esetek azonosítása és óvatos bevezetése jellemző, ugyanakkor, mint az MNB 2024. évi FinTech és digitalizációs jelentése rámutat: a hazai bankok kétharmada stratégiai és operatív szinten is számol az MI alkalmazási lehetőségeivel. Ezt követően szólt az eddigi uniós és hazai szabályozási kísérletekről, részletesebben elemezve a világszinten is úttörőnek számító AI Act pénzügyi fogyasztóvédelmi vonatkozású rendelkezéseit.

Farkas Jácint tudományos főmunkatárs (BGE PSZK Vezetés és Emberi Erőforrás Fejlesztés Tanszék) *Az MI-k védik meg a fogyasztókat vagy „Tőlük” kell (meg)védeni az embereket?* kérdéseit feszegetve rámutatott, hogy hermeneutikai csapdával állunk szemben a mesterséges intelligencia kapcsán. Az intelligencia fogalmát értelmezve leszögezte, hogy a „mesterséges intelligencia” nem talál ki új dolgokat, csak olyan inputokkal dolgozik, amelyekkel mi látjuk el őket, így tulajdonképpen nem tesznek mást, mint mintázatokat keresnek. Mindazonáltal, ha eléri az MI által generált információ mennyiség az ember által előállított mennyiséget, az emberi létezés elveszítheti lényegét. Az MI képes lehet olyan összefüggéseket találni, amelyeket mi nem, de a gondolkodás terepnumát nem szabad átengedni. A jogszabályok önmagukban nem fognak megfelelő védelmet garantálni, ha nem értjük meg az MI lényegét (ehhez szemléletes adalékként az előadó rámutatott, hogy Ausztráliában sokáig *mutated intelligence* volt a megnevezés). Csak ezt követően lehetne lehatárolni, hogy az MI mire használható, és mire nem. Az MI megszemélyesítése mindazonáltal – az előadó szerint – veszélyes irányt.

Toth Gedeon főiskolai tanár, szaktanácsadó (BGE KKK Marketing Tanszék) és Szabó Edit mb. oktató (BGE KKK Marketing Tanszék) *A mesterséges intelligencia fogyasztóvédelmi kérdései a reklámban* címmel tartott előadásukat azzal a felvetéssel kezdték, hogy a 2000-es évek elején a magyar marketingben gyakorlatilag senki nem költött internetes reklámokra, azonban mára ez az arány 50% fölé emelkedett, visszaszorítva a korábbi marketing felületeket (pl. televízió, rádió, nyomtatott sajtó). Ami a mesterséges intelligenciát illeti, az képes hiperperszonalizált reklámok előállítására, de a jogszabályi megfelelés ellenőrzéséhez is hasznos lehet. Mindazonáltal az MI önmagában sem jó, sem nem rossz; egy eszköz, amelyet etikusan kell használni. Ennek kulcsfontja, hogy a fogyasztót megfelelően tájékoztatni kell az MI alkalmazásáról. Az etikus használat elősegítése érdekében a marketingszakmai szövetségek fehér könyvet készítenek kiadni.

Szőke Andrea főiskolai docens (BGE KVIK Vendéglátás Tanszék) és Fekete-Frojimovics Zsófia tanszékvezető, egyetemi docens (BGE KVIK Vendéglátás Tanszék) a *Digitális hitelesség az influenzszerek világában: mi van, ha MI van?* kérdéseit elemezték. Az előadás kiindulópontjaként megállapítható, hogy 2024-ben globálisan a vállalatok 37%-a alkalmazott generatív MI-t marketingtevékenysége során. A hitelesség

témakörében az előadók kutatást végeztek a BGE hallgatói körében, amelynek során egy 9 kérdésből álló online kérdőív került összeállításra, amelyet ezidáig 83 fő töltött ki. A kérdőív tanulságait elemezve megállapítható, hogy a kitöltők a legnagyobb veszélynek a hitelesség elvesztését és a hamis információk terjesztését tartják. Az influenzaszerek általi MI-használatot akkor tartják elfogadhatónak, ha az átlátható, jelzik annak használatát. A szabályozás iránti igény erős, az azzal szemben támasztott elvárások között pedig leggyakrabban a csalás megelőzése, a személyiségi jogok és a biztonság védelme, a társadalmi hátrányok csökkentése és a fogyasztóvédelem szerepeltek.

Hámori Antal kutatócsoport-vezető, egyetemi docens (BGE KVIK Vendéglátás Tanszék) *A mesterséges intelligencia és a katolikus egyház – fogyasztóvédelmi nézőpontból* címmel tartotta meg előadását, rámutatva, hogy ez a „rendkívüli technológiai vívmány” a katolikus egyház érdeklődését is felkellette. Mindezt számtalan, a témában kiadott dokumentum megjelenése is mutatja.⁶ Az *Antiqua et nova* dikasztériumi dokumentum az addigi vonatkozó szentszéki, pápai megnyilatkozásoknak mintegy összefoglalását adja; etikai iránymutatásai szerint az emberi méltóságot tiszteletben tartandó, a tudományos és a technológiai fejlődésnek az emberi személy szolgálatára és a közjóra kell irányulnia, így a mesterséges intelligenciát is ezek érdekében kell felhasználni. Az egyház elismeri, hogy a technológia számtalan rosszat orvosolt, azonban különösen ellenzi azokat az alkalmazásokat, amelyek veszélyeztetik az élet szentségét, az emberi személy méltóságát, vagy akár a környezetet, továbbá rámutat, hogy a mesterséges intelligencia komoly kockázatokat, veszélyeket, csábításokat is jelent például a manipulált tartalom, a hamis információk tekintetében, amelyeknek a fogyasztók is ki vannak téve. A dokumentum záró gondolatai között hangsúlyozza, hogy a mesterséges intelligenciát – az emberi intelligencia kiegészítőjeként, nem pedig az emberi intelligencia gazdagságának „helyettesítőjeként” – csak eszközként szabad használni.⁷

Veres Zoltán – Hámori Antal
megbízott oktató, – egyetemi docens,
PPKE JÁK BGE KVIK

⁶ Ezek közül az előadó kiemelte az *Antiqua et nova* (AN) kezdetű, a mesterséges intelligencia és az emberi intelligencia kapcsolatáról szóló, Ferenc pápa által 2025. január 14. napján jóváhagyott közös dikasztériumi dokumentumot, Ferenc pápa üzenetét a béke 57. világnapjára (2024. január 1.) *A mesterséges intelligencia és a béke* címmel, valamint Ferenc pápa üzenetét a tömegtájékoztatás 58. világnapjára (2024. május 5.) *Mesterséges intelligencia és a szív bölcsessége: a teljesen emberi kommunikációért* címmel, továbbá Ferenc pápának a G7-ek ülésén, 2024. június 14-én a mesterséges intelligenciáról elhangzott, *Lenyűgöző és félelmetes eszköz* című beszédét, illetőleg Ferenc pápa üzenetét a 2024. évi teremtésvédelmi imapanpra (2024. szeptember 1.) *A teremtett világgal együtt remélj és cselekedj* címmel.

⁷ Ld. AN 112. pont.

SUMMA

FINDING OUR FIRST COMMEMORATIVE LAWS

Gábor Hollósi

Although the number of commemorative laws in our legal system is relatively high today, as they contain little or no binding rules of conduct, they raise the question of the concept of law with elementary force. Can there be a place for historical memory in the law, or does the legislator only burden the legal system with it unnecessarily? In my study, however, I do not deal with these exciting theoretical problems, but focus on the historical tradition of our commemorative laws, first and foremost looking for the first legislation that can be clearly described as commemorative. I will then present the world of our early commemorative laws of the first half of the nineteenth century, grouping the legislative products of the diets of the estates under four headings: the commemorative laws on recording decisions of ecclesiastical relevance, on foundations, on the founding of our great national institutions, and the birth of the “modern” commemorative law. Finally, I will explain my position on why I consider the legal materials I have presented to be commemorative, despite the differing opinions of authoritative authors, which also explains why I am attempting to draw a comprehensive and contrasty initial arc in the historical development of our commemorative law. My work is primarily based on the legal materials researched from our historical law database and their systematisation according to the above-mentioned aspects – in this article, the reader can get acquainted with the material of the first chapter of my book “Our Commemorative Laws”, which is being prepared and is planned to be published only in 2027, and can express his opinion on it before the manuscript of the book will be closed.

* * *

STRUGGLE FOR THE LATVIAN STATE LANGUAGE

Krisztián Manzinger

For Latvia, which restored its independence in 1991, language policy has been a matter of paramount importance for historical reasons. The basic pillars of the current language regulation are the Constitution and the State Language Law, but many sectoral laws also contain provisions, especially in the fields of citizenship,



education, media or name use. The primary goal of the legislator has been to change the linguistic landscape of Latvia, which shifted towards unbalanced, Russian-dominated bilingualism during the decades of the Soviet occupation, in a way that ensures the dominance of the Latvian language and pushes not the Latvian majority but the national minorities towards bilingualism. The first State Language Law of 1989, passed before the restoration of the independence, was based on the Latvian-Russian bilingualism of the state, in which residents would have had the right to decide in which language they would address institutions. However, after the restoration of the independence, but before the law entered into full force in 1992, the legislator amended the regulation in a way that declared Latvian the exclusive official language. The 1998 constitutional amendments and the new 1999 State Language Law completely prohibited the use of languages other than Latvian not only within public institutions, but also in the communication between residents and public institutions – except for a few enlisted, primarily emergency situations – and significantly restricted the free language use in the private sphere. Despite this, the legislator’s goal of ensuring the dominance of the Latvian language has still not been fully achieved, leading to continued legislation on sectoral restrictions in recent decades, criticized by those concerned in Latvia and international forums protecting minority rights.

* * *

ARTIFICIAL INTELLIGENCE LITERACY

Mónika Andok

Media literacy and AI literacy are essential skills of the 21st century. They are indispensable in education, the labor market, and digital citizenship. Therefore, it is of utmost importance to define the skills precisely, to define the framework of use, and to develop measurement tools and to make them known to the professional audience. The study focuses on three main areas in this regard. First, it presents the relationship between media literacy and AI literacy. Then, it introduces the AI literacy framework in higher education. Finally, it presents tools suitable for measuring AI literacy. Accordingly, the study first presents the UNESCO Digital Skills Framework. After media literacy, the study discusses the issue of AI literacy. AI literacy has been interpreted and defined in various ways. Ng and colleagues identified four main areas: knowledge and understanding of AI, the ability to use and apply AI, the ability to evaluate and create AI, and ethical considerations. The paper then presents the UNESCO AI Literacy for Students framework, which outlines 12 competencies in four dimensions: Human-Centered Thinking, Ethics in AI, AI Techniques and Applications, and AI System Design. These competencies span three levels of development: Understanding, Applying, and Creating. In the third major unit of the paper, we present the available measurement tools from the perspective of AI Literacy

* * *

THE FREEDOM OF EXPRESSION OF ARTIFICIAL INTELLIGENCE AND ITS CONTENT MODERATION IN THE DSA AND THE AI ACT

Gergely Gosztonyi – Dorina Gyetván – Andrea Kovács

After a brief introduction to AI, the paper will discuss whether AI has freedom of expression or whether it is only available to the companies that own it. This is important in the context of whether AI can be considered a separate legal actor in the context of content creation in social media, which is the second main issue of the paper. Here, negative and positive cases of content moderation are presented, i.e. whether the full spectrum of content hosting in social media can be entrusted to AI as a separate actor with respect to liability, or whether it requires some form of human control, direction or supervision.

* * *

TRANSPARENCY CHALLENGES OF AI-ASSISTED CONTENT MODERATION UNDER THE DIGITAL SERVICES ACT

Balázs Hohmann

The study, titled “Transparency Challenges of AI-Assisted Content Moderation Under the Digital Services Act,” critically examines the transparency challenges inherent in AI-assisted content moderation within the regulatory framework of the Digital Services Act (DSA). As AI technologies have become integral to content moderation, enhancing efficiency and scalability, they simultaneously introduce complex legal and ethical dilemmas. The DSA aims to strengthen platform transparency, safeguard fundamental rights, and ensure the accountability of moderation processes. A pivotal question thus emerges: how can these regulatory objectives be effectively realized in an ecosystem where AI-driven moderation mechanisms predominate?

This study underscores that the transparency obligations imposed by the DSA, particularly in relation to algorithmic decision-making, present unique compliance challenges for AI-based moderation systems. It conducts a comprehensive analysis of the relevant DSA provisions, the compliance strategies adopted by online platforms, and the enforcement proceedings initiated by the European Commission under Article 66, with a particular emphasis on AI governance.

The findings indicate that while the DSA constitutes a significant advancement in regulatory transparency, ensuring full transparency in AI-assisted moderation remains a formidable challenge. Addressing these persistent gaps may necessitate a critical reassessment of the regulatory and operational frameworks governing AI-driven content moderation.

ALGORITHMIC BIAS AS A LEGAL POLEMIC

Gergely Ferenc Lendvai

The paper explores algorithmic bias as a legal dilemma, presenting the conceptual basis of the phenomenon, its effects and regulatory challenges. The fragmented regulatory environment in the US and the EU's strict, user-centric framework, such as the GDPR, DSA and AI regulations, are presented. As a result of this paper, we highlight the key role of transparency and accountability mechanisms in mitigating bias and emphasise the urgent need for coordinated global regulatory efforts. The study also points out that existing enforcement mechanisms are often insufficient for effective implementation. In particular, the problem of 'black box' algorithm design persists, which hinders the identification and remediation of biases, and regulatory frameworks show significant inequalities between different groups. In this paper, we conclude that without a globally harmonized regulatory framework, existing efforts cannot effectively address the challenges of AI, in particular bias. In the paper, we urge global cooperation, which we believe is essential to respond appropriately to technological developments.

* * *

ALGORITHMS AND TRANSPARENCY

*the Legal Regulation of Recommender Systems
on Online Platforms in the European Union*

Kitti Mezei

The increasing dominance of digital services has significantly reshaped online information flows, content consumption habits, and platform operations. The European Union's Digital Services Act (DSA) aims to regulate the legal environment of online services, ensuring user rights, transparency, and a balanced digital market. Recommendation systems, which influence content accessibility and visibility, play a crucial role in this regulatory framework. Their algorithmic decision-making processes impact public opinion, market competition, and fundamental rights, prompting the EU to enforce transparency and accountability in ranking mechanisms, targeted recommendations, and automated decision-making. This study explores the DSA's provisions on recommendation systems, analyzing their practical implications and assessing their legal and ethical impact. It examines platform obligations, transparency requirements, and redress mechanisms that protect user rights. Ethical concerns, such as algorithmic bias, filter bubbles, and manipulation risks, are discussed, along with the tension between transparency and excessive information overload. By critically evaluating the DSA's approach, the study highlights key challenges in implementation

and enforcement. It argues that while transparency mandates are a step forward, real user control remains limited. Without clear regulatory guidelines and oversight, platform compliance may lead to mere “transparency washing” rather than substantive change.

* * *

IS ARTIFICIAL INTELLIGENCE OPENING A NEW ERA OF DISINFORMATION?

*Understanding the role of AI using the disinformation chain
as a new theoretical framework*

Róbert Pintér

This paper examines the relationship between artificial intelligence (AI) and disinformation and seeks to answer whether AI is opening a new era in disinformation. The author conducts a systematic investigation, introducing disinformation chain as a new concept. Using this theoretical framework, the paper describes the motivation, creation, dissemination, detection, reception and wider social impact of disinformation. It also describes the application of AI in these steps. The role of AI is particularly significant in the dissemination and social impact of disinformation, while the emergence of generative AI has not proved to be decisive in the creation of disinformation, contrary to prior expectations. If we can talk about a new era, it is more related to social media algorithms and participatory propaganda with the active involvement of people, which has led to the emergence of a post-truth era since 2016. The article also proposes a set of actions for effective measures against disinformation along the disinformation chain. Finally, the author concludes by expressing his belief that the statement ‘we cannot believe our eyes and ears’ crystallises not as a fatalistic creed but as a healthy scepticism, since no society can function without human trust.

* * *

A GUIDE TO DRAFT A SCHOOL (CYBER)BULLYING POLICY

Tamás Pongó

The phenomenon of bullying and cyberbullying unfortunately enmeshes the Hungarian public education system, and many students, parents and teachers suffer from the trauma of bullying. The problem can only be addressed through an interdisciplinary approach, which requires the tools of psychology and pedagogy as well as the tools of law. The aim of this paper is to develop a model anti-bullying policy that provides a

common minimum set of guarantees and a starting point for primary and secondary schools to develop their own independent policy. In the annex to the study, a model policy has been drafted, the elements of which fit organically into the existing legal framework and complement it with the scientific background and elements of bullying. By way of example, in addition to addressing the issue of personal and territorial scope, the policy defines bullying and cyberbullying behavior, as well as the retaliation that often occurs in the course of this phenomenon. It also offers a recommended set of procedures, which, if fine-tuned, can be used to deal with incidents of bullying in an appropriate manner.

* * *

THE SOCIAL EFFECTS OF ARTIFICIAL INTELLIGENCE-BASED CONTENT FILTERING ALGORITHMS, WITH SPECIAL REGARD TO SOCIAL MEDIA PLATFORMS

István Üveges

Online communication spaces, especially social media platforms, have become the dominant means of information flow, but they are increasingly exposed to the spread of extremist ideologies, illegal and sometimes disturbing content. Such platforms are increasingly using AI-based algorithms to filter and moderate content to address these challenges. The purpose of this study is to provide a comprehensive picture of the reasons for the spread of automated content moderation, its technical solutions, and its social effects. The study discusses in detail the legal environment that allowed these systems to take off, presents the most used technical methods, and analyzes how they affect freedom of speech, social discourse, and political polarization. The results highlight that although artificial intelligence can significantly increase the effectiveness of moderation, its use involves serious challenges and risks that affect democratic values and the functioning of social media. The study also makes recommendations on the importance of regulation and human oversight to ensure the protection of free speech and the preservation of social diversity in the online space.